ESTIMATING THE RELATIVE TREATMENT EFFECTS OF NATURAL CLUSTERS OF ADOLESCENT SUBSTANCE ABUSE TREATMENT SERVICES: COMBINING LATENT CLASS ANALYSIS AND PROPENSITY SCORE METHODS

by

Megan S. Schuler

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

December 12, 2013

ABSTRACT

Objectives: The motivating substantive aim of this dissertation was to identify common clusters of drug treatment services that adolescents receive *in practice* that are effective in terms of improving substance use outcomes. We first identified clusters of drug treatment services that adolescents in outpatient treatment report receiving, as well as examined factors associated with each class of treatment services (Chapter 2). Our statistical approach for estimating the effect of treatment service classes on outcomes was latent class regression with a distal outcome; we review various statistical methods for implementing latent class regression with a distal outcome in Chapter 3. Addressing potential confounding arising from baseline differences among youth receiving different classes of treatment services was a key concern; Chapter 4 describes emerging methods to address confounding in the context of latent class regression with a distal outcome, highlighting the challenges that arise when the treatment of interest is a latent variable.

Methods: Chapters 2 and 4 used data on 5,527 adolescents receiving drug treatment services through treatment providers funded through the Substance Abuse and Mental Health Services Administration's Center for Substance Abuse Treatment. Latent class analysis was used to identify classes of substance use treatment services reported by youth. A simulation study to compare 5 statistical methods for latent class regression with a distal outcome was performed in Chapter 3. An additional simulation study to compare 3 methods for addressing confounding in this context was performed in Chapter 4; these methods were also applied to our adolescent data.

Results: Distinct classes of outpatient treatment services received by adolescents were empirically identified using latent class analysis; youth receiving different classes of treatment services were found to be significantly different on numerous baseline characteristics. Statistical performance varied notably across methods for latent class regression with a distal outcome. Finally, failing to account for potential confounding in this setting can lead to significantly biased estimates of the association between the latent class and the distal outcome; the 1-step method we examined performed particularly well in terms of reducing bias.

Conclusions: Emerging methods for modeling the treatment of interest as a latent variable are quite relevant for social and behavior researchers. However, like studies with fully observed variables, care must be taken to address potential confounding; future work should continue to develop methods to address confounding in this context.

THESIS ADVISOR:	Elizabeth Stuart, Ph.D.
THESIS READERS:	Rosa Crum, M.D., M.H.S. (Chair)
	Jeannie-Marie Leoutsakos, Ph.D.
	Elizabeth Letourneau, Ph.D.
ALTERNATES:	Kitty Chan, Ph.D.
	Ramin Mojtabai, M.D., Ph.D., M.P.H.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Liz Stuart for all of her wisdom, guidance, and support. It has been an absolute privilege to work with her and she remains an ongoing source of professional inspiration to me. Her optimism and endless new ideas were profoundly helpful, especially during the doldrums of the grad school experience. I value all of the wonderful opportunities she facilitated for me and all of the ways in which she pushed me intellectually; these experiences truly made my time at Hopkins so enriching and meaningful. I look forward to ongoing collaborations with her in the future.

I also would like to thank Beth Ann Griffin for her role in the genesis of this project. This dissertation work grew out of a project I worked on with Beth Ann while a Summer Associate with the RAND Corporation, and she was quite instrumental in conceptualizing this work. I am grateful to her and RAND for allowing me ongoing access to the data that was central to this dissertation.

I would also like to thank my committee members Rosa Crum, Jeannie Leoutsakos, and Elizabeth Letourneau for their contributions. I really value the thoughtfulness that each of them has brought to this project; their insights and critiques have assuredly strengthened this work. Ramin Mojtabai and Carl Latkin also both provided very helpful feedback during earlier stages of this work.

Finally, I am incredibly grateful for the love and support from friends and family through my entire academic journey, and during especially the past four years. Few emotions are best experienced alone; the complex range of emotions experienced en route to a completed dissertation proves no exception. I am particularly indebted to my wonderful parents, who have been my unwavering #1 fans from the beginning and to Anthony who, over many years, truly taught me what it means to never give up.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. Introduction

1.1 Problem statement

Substance use problems among adolescents are of notable public health concern, given the significant health and social consequences of substance use during adolescence as well as the elevated risk for subsequent substance problems during adulthood. In recent national surveys, 16.8% of youth report past month illicit drug use and 6.9% of youth ages 12 to 17 meet DSM-IV criteria for a substance use disorder (Johnston et al., 2013; SAMHSA, 2012). Identifying effective treatment programs and improving access to treatment is essential, since only around 10% of youth needing treatment receive any formal substance treatment (SAMHSA, 2012).

There has been concerted efforts to identify effective substance treatment programs for adolescents, which have resulted in the classification of seventeen adolescent substance disorder treatment programs as "evidence based" as defined by the Substance Abuse and Mental Health Services Administration's (SAMHSA) National Registry of Evidence-based Programs and Practices (NREPP) (SAMHSA, 2012). However, evaluation of adolescent substance treatment programs has significantly lagged behind evaluation of adult substance treatment programs, leaving many unanswered questions about what treatment services are optimal for youth. One limitation to many existing studies is that they classify treatment groups based on the program that youth were *randomized to* (in an RCT) or were *enrolled in* (in an observational study). Although such efficacy studies are necessary, there has been little work to understand the effectiveness of treatment services that youth receive *in practice.*

Program enrollment may not reflect services actually received for a variety of reasons. Youth may receive additional services due to concurrent enrollment in treatment programs; for example, individuals may complement clinical treatment with participation in self-help groups (Kelly & Myers, 2007). Similarly, adolescents involved in the criminal justice system may be mandated to enroll in a treatment program (through public or private treatment providers) as a condition of parole or probation, while simultaneously receiving services through the criminal justice system, such as case management services or drug screening. On the other hand, adolescents may receive fewer services than expected in a given program due to noncompliance or dropout, both of which are significant problems in the treatment of adolescent substance use and may reflect a youth's lack of intrinsic motivation.

Thus, one objective of this work was to empirically identify and describe common clusters of treatment services that youth receive through outpatient drug treatment. Data for this study came from adolescent treatment providers throughout the US funded through the Substance Abuse and Mental Health Services Administration's Center for Substance Abuse Treatment (CSAT). Youth in this study were assessed using the Global Appraisal of Individual Needs (GAIN) survey, which included items about whether youth received specific drug treatment services. Based on these items, we performed latent class analysis to identify classes of treatment services youth reported receiving. Additionally, we performed latent variable regression to identify baseline characteristics, including demographics, substance use, and justice system involvement, that were associated with membership in the latent classes we identified.

Additionally, we were interested in estimating the causal effect of these latent classes of treatment services in substance use outcomes. This effect can be estimated using latent variable regression with distal outcomes, in which a given substance use outcome (a traditional, fully observed variable) is regressed on treatment class (a latent variable). Methodologically, there are several approaches to estimating latent variable regression models, broadly characterized as *1-step methods*, which jointly estimate the measurement model (i.e., the latent class model) and the structural model (i.e., the link between the latent variable and outcome), and *3-step methods*, which sequentially estimate the measurement model and then the structural model. One-step methods are preferable in terms of statistical efficiency and in that they yield unbiased parameter estimates under correct model specification, yet 1-step methods may not be computationally or conceptually appropriate for all applications. Three-step methods, as classically implemented, are known to yield significantly biased estimates. Although corrected 3-step methods that perform relatively similarly to 1-step methods have been proposed in recent years, these methods have not been widely disseminated or adopted. Thus, the second objective of this project is to provide a comprehensive overview for applied researchers of 1-step, classical 3-step, and corrected 3-step methods available for latent variable regression.

Furthermore, an additional complication when estimating the effect of latent treatment classes on substance use outcomes is the potential for confounding. Our data on adolescent substance use comes from an observational design, and thus it is likely that the adolescents who receive particular clusters of treatment services are quite different from those who receive different clusters of services. For example, the services an adolescent receive may be associated with his or her baseline substance use severity. Analysis methods that do not account for baseline differences between treatment groups may conflate these preexisting differences with the true treatment effects. Although there are many statistical methods, including propensity score methods, to address confounding in settings when all variables are fully observed, methods to address confounding in the presence of latent variables (especially a latent treatment, as we have here) are quite recent. Thus, a third objective of this project is to investigate current methods to address confounding in the context of latent variable regression. One method, known as Latent Class Causal Analysis (Kang & Schafer, 2010; Schafer & Kang, 2013), is a 1-step approach that allows joint estimation of a model that includes a latent treatment variable, confounders, and a distal outcome. As in the simpler case of latent variable regression without confounding, this 1-step method is expected to perform quite well, yet may not always be a feasible approach given the computational complexity of estimating the joint model or in the case of distal outcomes, may not always be conceptually appropriate, since 1-step methods allow the distal outcome to influence class formation. Additionally, 1-step methods are not compatible with propensity score methods, which may offer statistical advantageous in some settings and may be more robust to model misspecifiation. Therefore, we also investigate extensions of 3-step approaches that use propensity score methods to account for confounding. In brief, after estimating the measurement model, these approaches estimate propensity scores based on the measurement model, and then incorporate propensity scores when estimating the structural model. We investigate several methods for latent variable regression in the presence of confounding on fully simulated data in order to assess and compare the statistical performance of both methods. Additionally, we apply these methods to our adolescent substance use dataset in order to estimate the relative effectiveness of natural classes of treatment services on substance use frequency and consequences. Although our motivation example of interest comes from the field of substance use research, the statistical methods discussed in this dissertation have widely applicable in the fields of social, behavioral and health research.

1.2 Methodological Background

Latent Variable Modeling

Latent variable modeling was motivated by recognition that many constructs in the social and behavioral sciences are not fully observable (Clogg, 1995). For example, depression or self-esteem are constructs that investigators widely agree exist, yet are not directly measurable. When studying such constructs, symptom checklists or surveys consisting of multiple items are often used. Although these items reflect the underlying latent construct, they do not perfectly measure it, given the inherent unobservable nature of a latent variable. This discordance between the underlying true construct and the observed (i.e., "manifested") indicators is known as *measurement error.* Statistical methods known as latent variable modeling were developed in order to appropriately account for the measurement error arising from the study of latent variables.

One widely used latent variable model is latent class analysis (Clogg, 1995; Collins & Lanza, 2010; Goodman, 1974; Habeman, 1979; Hagenaars & McCutcheon, 2002; Lazersfeld &

Henry, 1968). Latent class analysis (LCA) models a categorical latent variable based on multiple, observed indicators $U_1, U_2, \ldots, U_j$, such that each individual belongs to exactly one of C latent classes, denoted $C_1, C_2, \ldots, C_C$. The latent classes are defined by response patterns among the indicator items – for example, one class may have a low propensity for endorsing all items while another class is defined by high propensities for endorsing all items.

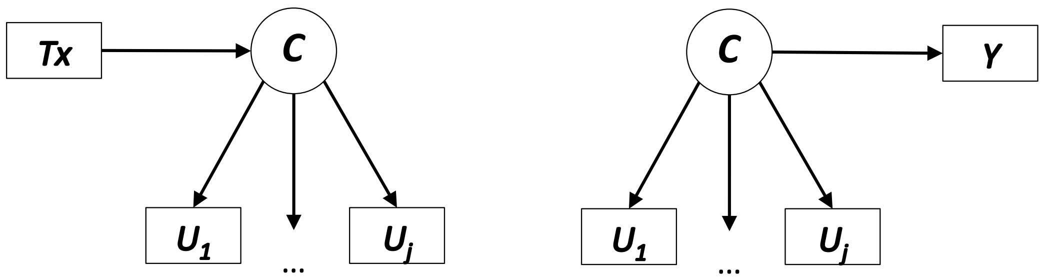
The two parameters of primary interest in a latent class analysis are the posterior probabilities of latent class membership and the conditional item probabilities. Posterior probabilities of class membership represent the probability that an individual with an observed pattern $\boldsymbol{u}$ on the indicators will belong to latent class c, and are denoted $\gamma_{c|u} = \Pr(C = c|\boldsymbol{U} = \boldsymbol{u})$. The conditional item probabilities represent the probability that an individual in latent class c will endorse a given indicator U_j. Formally, assuming binary indicators, the conditional item probabilities are defined as $\rho_j = \Pr(U_j = 1 \mid C = c)$. The conditional probabilities are used to interpret the meaning of each latent class.

Latent class analysis requires the following assumptions. First, latent classes are conceptualized as homogenous, in the sense that all individuals within a given latent class are estimated to have the same distribution with respect to the latent indicators. Also, individuals are assumed to be independent, such that a given individual's latent class membership does not affect another individual's. Finally, LCA assumes *local independence*, which states that the indicators $U_i, \ldots U_j$ are mutually independent after conditioning on latent class membership.

Latent Variable Regression

While latent class analysis is purely a measurement model, latent class regression introduces a structural component that allows the estimation of the association of latent class and an auxiliary variable. Generally, this auxiliary variable may either be observed or latent; in this project we focus only on latent variable regression in which the auxiliary variable is observed. We consider two types of regression models: (1) latent class regressed on a predictive covariates and (2) a distal outcome regressed on latent class (see Figure 1). Latent class regression typically assumes that the auxiliary variable does not have any direct effect on the indicators $U_i, \ldots U_j$, but rather is only associated with the indicators via class membership.

Figure 1.1. Two classes of latent variable regression models considered in this project.



Panel A. Latent class regression with predictive treatment

Panel B. Latent class regression with distal outcome

Confounding in Observational Studies

One of the major challenges in observational studies is obtaining valid (i.e., unbiased) estimates of treatment effects in the absence of directly comparable treatment groups. Estimating unbiased treatment effects is more straightforward in randomized controlled trials (RCTs) since randomization creates treatment groups that are equivalent ("balanced") with regard to both unobserved and observed variables. Under randomization, the observed differences in outcomes between treatment groups may be attributed to a treatment effect, since groups only differ in terms of treatment status. In contrast, in observational studies there may be systematic pre-existing differences between individuals in different treatment groups, a phenomenon referred to as selection bias. For example, in the context of substance use treatment, factors such as substance use severity, socioeconomic status, access to care, healthcare provider experience and personal preference may influence the types of treatment services that an individual receives. Statistical analysis of observational studies must carefully disentangle selection bias from true treatment effects.

Propensity Score Methods

In Aim 3 of this project, we explore the potential for incorporating propensity score methods with latent variable regression in order to address confounding. Propensity score methods are a standard statistical method for addressing the selection bias in an observational study (Rosenbaum & Rubin, 1983). Consider the traditional binary treatment case, in which an individual either receives the treatment or control condition. Formally, the *propensity score* is defined as the probability that an individual received the treatment, conditional on the individual's observed covariates. Specifically, $p(x) = \Pr(T_i = t | X_i = x)$ where *T* denotes treatment status (0 or 1) and *X* denotes a vector of observed covariates. More generally, the propensity score can be extended to cases in which there are multiple treatment groups. Imbens (2007) refers to this as the *generalized propensity score*, defined as $p(t, x) = \Pr(T_i = t | X_i = x)$ where $t \in \mathcal{T}$.

In an RCT employing simple randomization all individuals within a given treatment group have the same likelihood of receiving their assigned treatment (conditional on any covariates that were used in the randomization). However, in an observational study, individuals within a given treatment group might have varying likelihoods of receiving their observed treatment. As detailed below, propensity score methods control for each individual's propensity score, thereby addressing the problem of selection by creating statistically comparable treatment groups. Note that propensity score methods can only create balanced treatment groups with respect to observed covariates, whereas randomization creates treatment groups that are balanced with respect to both observed and unobserved covariates.

Propensity score methods are preferable to regression covariate adjustment for several reasons. First, propensity score methods do not necessarily rely on the parametric modeling assumptions required by regression adjustment (Ho et al., 2007). Additionally, propensity score methods avoid potential bias that arises from extrapolating beyond observed data in traditional regression models when the treatment groups have little overlap in terms of covariates (Stuart, 2010). Furthermore, propensity scores are an effective dimension reduction technique when there are a substantial number of baseline covariates to adjust for (Rosenbaum & Rubin, 1984). Finally, as advocated by Rubin, it is philosophically cleaner to separate the analytic step of controlling for confounding from the step of implementing the final structural model (Rubin, 2001). Separation prevents potential bias that may arise from adjusting for covariates solely because they favorably influence the treatment effect estimates.

Although there are numerous approaches to modeling propensity scores, a common method is to use logistic regression to model the probability of receiving the treatment as a function of observed covariates; nonparametric methods may also be used (McCaffrey et al., 2004). The primary methods for incorporating propensity scores in the final analysis are matching, subclassification, and weighting (Stuart, 2010).

1.3 Overview of specific aims and hypotheses

In Aim 1, we identify underlying latent classes of substance abuse treatment services among a large sample of adolescents receiving outpatient substance use treatment. Youth were assessed with the Global Appraisal of Individual Needs (GAIN) survey; the GAIN includes the Treatment Received Scale (TxRS) which asks whether youth received specific drug treatment services. Twelve items from the TxRS form the basis for latent class analysis.

Aim 1A: To identify latent classes of treatment services received by adolescents enrolled in outpatient substance use treatment programs using latent class analysis.

Aim 1B: To identify baseline characteristics, including demographics, substance use, and juvenile justice system involvement, that are associated with membership in the various classes of treatment services using latent class regression.

In Aim 2, we undertake a comprehensive review of 1-step, classical 3-step, and corrected 3-step methods for latent variable regression. This review is aimed at applied researchers and encourages adoption of 1-step and corrected 3-step methods, given that classical 3-step methods are still used in practice despite their known limitations. We compare the performance of various 1-step, classical 3-step, and corrected 3-step methods using fully simulated data.

Aim 2A: To provide a review of 1-step, classical 3-step, and corrected 3-step methods for latent variable regression for applied researchers.

Aim 2B: To compare the statistical performance of 1-step, classical 3-step, and corrected 3-step methods for estimating the association between latent class and a distal outcome, using simulated data.

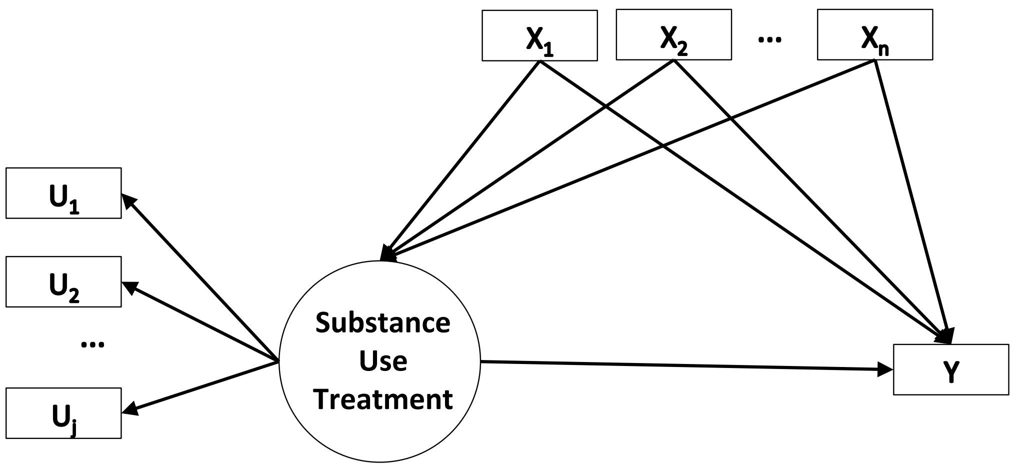
In Aim 3, we focus on latent variable regression in the presence of confounding. Given that latent variable regression is often performed on data arising from observational studies, we highlight the importance of recognizing and addressing potential confounding. Given that this is a relatively new area of statistical development, we review the few current approaches and propose extending 3-step methods by incorporating propensity score methods. We compare several methods using simulated data and then apply these methods to our adolescent substance use data.

Aim 3A: To compare the statistical performance of 1-step and proposed 3-step methods for estimating the association between latent class and a distal outcome, in the presence of confounding, using simulated data.

Aim 3B: Apply the 1-step and 3-step methods compared in Aim 3A to data from adolescents in outpatient substance use treatment, in order to estimate the causal effect of latent classes of treatment services on substance use frequency and substance use problems.

Figure 1.2 depicts the statistical model for Aim 3. Substance use treatment services will be modeled as a latent variable, defined by indicators $U_1, U_2, \ldots, U_m$. Outcomes of interest, namely substance use frequency and substance use problems are denoted Y_1 and Y_2, and potential confounders are denoted $X_1, X_2, \ldots, X_n$.

Figure 1.2. Statistical model for Aim 3.



The structure of this work is as follows. Chapters 2 – 4 discuss Aims 1, 2, and 3 respectively. Chapter 5 concludes by providing a synthesis of the findings of the three Aims, discusses strengths and limitations of the work, and outlines future research directions.

CHAPTER 2. Common clusters of drug treatment services received by adolescents: A latent class analysis

2.1 Abstract

Background: Although numerous evaluation studies have been conducted to compare adolescent substance use treatment programs, previous studies have generally assessed the effect of programs to which youth *were randomized to* or *were enrolled in*, rather than the effect of the services they *truly received in practice*. It is imperative that we understand current substance use treatment trends for youth, since identifying and describing patterns in service provision is a fundamental first step in determining what services work most effectively in practice.

Methods: This study included 5,527 adolescents who were receiving outpatient drug treatment services from treatment providers throughout the US who were funded by the Substance Abuse and Mental Health Service Administration's Center for Substance Abuse Treatment. All youth were assessed with the Global Appraisal of Individual Needs (GAIN). Using 12 items from the Treatment Received Scale that spanned the domains of individual-focused, family-based, and case management services, latent class analysis was performed to identify common clusters (classes) of treatment services youth reported receiving. Latent class regression identified youth characteristics associated with each class of services.

Results: Four latent classes were identified and are described as: Class 1: Low Service Utilization (12% of youth); Class 2: Individual-Focused Services (39%); Class 3: Individual- and Family-Focused Services (38%); and Class 4: Multiple Services (11%). Latent class regression identified significant differences across classes with regard to demographics, factors related to substance use, and justice system involvement.

Conclusions: This study demonstrates that, among a large population of adolescents receiving outpatient services through various treatment programs in the US, distinct groups of youth can be empirically identified based on the services they received.

2.2 Introduction

To date, most evaluation studies on adolescent drug treatment have been in the context of randomized trials and have investigated the effectiveness of the treatment services *to which youth were randomized*. Existing observational evaluation studies typically investigate the effectiveness of treatment services *in which youth were enrolled.* A separate, but highly relevant, question is what is the effectiveness of treatment services *that youth actually receive in practice*. The first step in answering this question is to identify and describe common clusters of treatment services that receive in outpatient treatment settings, which is the objective of this paper. Although this approach requires appropriately assessing services received by youth, classifying youth with regard to service provision allows researchers to identify commonalities across programs as well as to ultimately examine the effectiveness of current practices in adolescent drug treatment.

Although there is great heterogeneity across treatment programs, due to different target populations, treatment settings, and treatment philosophies, there are also many commonalities. One core approach is individual-focused therapy, which emphasizes the youth's central role in affecting change in his or her behavior. The most commonly used theoretical approaches are cognitive behavioral therapy (CBT), motivational interviewing (MI) / motivational enhancement therapy (MET), and contingency management (CM) (Winters et al., 2011). CBT espouses that behavior can be changed by modifying and re-scripting cognitions; therapy focuses on building skills such as managing unwanted drug-related thoughts, improving drug refusal skills, and identifying substance-free activities (Akers et al., 1979). MI and MET seek to increase an individual's intrinsic motivation for change (Miller & Rollnick, 1991). CM is an approach in which contingencies, typically positive reinforcement, are associated with a desired behavior, such as a negative drug screen (Higgins et al., 2008; Petry & Simic, 2002).

Another common and long-standing approach is family-based therapy, which views adolescent behavior within the context of his or her environment, recognizing that the environment contributes both risk and protective factors. Adolescent substance use has been found to be associated with family factors including parental psychopathology, poor parental monitoring, marital conflict, and low family cohesion (Diamond & Josephson, 2005; Rowe & Liddle, 2003); family-based interventions seek to reduce such risk factors and improve protective factors, such as improving family communication. Two prominent evidence-based models are Multidimensional Family Therapy (MDFT; Liddle, 1999, 2004) and Multisystemic Therapy (MST; Henggeler et al., 1998, 2009).

Case management services are additional support services provided to youth who are receiving or recently completed treatment in order to improve service continuity and outcomes (Vanderplasschen et al., 2004). Since youth with substance use problems often face significant challenges in other life domains, case managers help youth work toward substance use recovery, as well as more broadly improving their vocational status, mental and physical health, social support, and family relationships (Godley et al., 1994). A prominent model is Assertive Continuing Care (Godley et al., 1994; 2002; 2007) which emphasizes a positive, non-confrontational approach to supporting youth in developing healthy family and peer relationships, improving school performance, and complying with parole/probation requirements.

An important complication to evaluation studies is that program enrollment often does not truly reflect the treatment services that youth receive. Factors such as noncompliance, treatment fidelity, and dropout can all result in a youth receiving fewer services than expected. The majority of youth are referred to treatment by the juvenile justice system or the education system (Godley & Godley, 2011); a lack of intrinsic motivation can lead to noncompliance with treatment protocol. Dropout also significantly impacts service provision, since nearly half of adolescents drop out of drug treatment by 6 weeks (Godley et al., 2005; Riley et al., 2012). Additionally, program fidelity can decline when programs are taken to scale (Hallfors et al., 2006), such that program descriptions may not accurately represent the services or quality of services implemented in practice. On the other hand, adolescents may receive services through multiple entities simultaneously, so a given treatment program may not fully capture the totality of services. Youth in the juvenile justice system may receive additional case management services through their parole/probation officer or as a condition of drug court. Also, participation in self-help groups is often recommended as adjunct therapy (Kelly & Myers, 2007). Overall, although many factors that impact service provision have been identified, relatively little is known about the combinations of services that youth actually receive in typical outpatient settings.

Using a national database from adolescent treatment providers, this study empirically identifies common clusters (classes) of treatment services reported by youth in outpatient drug treatment using latent class analysis. We then examine characteristics that are associated with these classes of treatment services using latent class regression. This study is unique in both defining treatment based on services youth report receiving rather than based on program enrollment and in using latent class analysis to identify common clusters of services among youth served by various treatment programs.

2.3 Methods

Participants (n=5,527) were obtained from the Substance Abuse and Mental Health Administration's Center for Substance Abuse Treatment's (CSAT) 2007 database. Youth were enrolled in one of 9 CSAT-funded treatment programs: the Effective Adolescent Treatment (EAT) program that supported MET/CBT-5 implementation (Dennis et al., 2004; Melchior et al., 2007; SAMHSA, 2003); the Cannabis Youth Treatment experiment (Dennis et al., 2004; Diamond et al., 2002) which randomized youth to MET, CBT, family-based therapy, or MDFT, or Adolescent Community Reinforcement Approach (ACRA); the Adolescent Treatment Models program (Dennis et al., 2003) providing community-based care; the Adolescent Residential Treatment (SAMHSA, 2002) providing continuing care for discharged youth; the Strengthening

Communities' Youth program (Dennis et al., 2008) aimed at building partnerships among community, school-based and juvenile justice treatment services for early intervention, outpatient and intensive outpatient programs; the Targeted Capacity Expansion program (Wilson et al., 2005) providing intensive outpatient and inpatient services; the Young Offenders Reentry Program (SAMHSA, 2004) providing services to youth re-entering the community; the Family and Juvenile Treatment Drug Court program (SAMHSA, 2005) providing comprehensive services through drug courts; and the Assertive Adolescent and Family Treatment program (Godley et al., 2007) promoting family-centered services.

This study was restricted to youth ages 12-18 who reported receiving no inpatient or residential drug treatment services during the study period (baseline to 3 months) and who were community-dwelling at baseline. Written informed consent from parents and assent from the adolescents were obtained, and institutional review boards at each site approved study protocol.

The primary instrument in this study is the Global Appraisal of Individual Needs (GAIN; Dennis, 1999). The GAIN assesses demographics, substance use and substance use treatment, risk behaviors, mental and physical health, legal status, environment risk factors, and education/vocation status. All GAIN items are based on youth self-report. Reliability studies conducted by Dennis et al. (1999, 2000, 2002) reported that the majority of the GAIN indices have a Cronbach's α greater than 0.85.

Substance treatment services received by youth from baseline to 3 months were assessed with the GAIN's Treatment Received Scale (TxRS). This scale includes 20 items about specific substance treatment services and is comprised of three subscales that measure Direct, Family, and External services (Dennis et al., 2010). Twelve items from the TxRS were used as indicators in the latent class analysis, due to redundancy among remaining items or infrequent responses (<5%). Specifically, four items from the Direct subscale were used to measure individual-focused services, four items from the Family subscale were used to assess family-based services and four items from the External subscale were used to measure case-management services.

Demographic variables included age, sex and race/ethnicity (categorized as White, Black, Hispanic, and Other). Baseline substance use was assessed by indicators for daily use, prior treatment, and recognition of substance problems, as well as the Substance Frequency Scale (past 90 days), Substance Problems Scale (past month), Substance Dependence Scale (past year), and Treatment Motivation Index. Legal status was assessed by indicators for justice system involvement, any arrests, any days in a controlled environment (all with respect to past 90 days), as well as number of days involved with illegal activities (past 90 days), and the Crime Violence Scale. Mental health was assessed by indicators for suicidal thoughts, and mental health treatment as well as the number of days affected by emotional problems (past 90 days), the Internal Mental Distress Scale, Behavioral Complexity Scale, and Problem Orientation Scale. Additional covariates included Living Environmental Risk Scale, Social Environmental Risk Scale, and indicators for current school attendance and employment status.

Latent class analysis was performed to identify common clusters of drug treatment services received. To determine the optimal number of latent classes, models with 1 to 6 classes were estimated and multiple fit statistics were assessed (Nylund et al., 2007), including entropy (classification error), Akaike Information Criterion (AIC; Akaike, 1974), Bayesian Information Criterion (BIC; Schwartz, 1978), and sample size adjusted BIC (a-BOC; Sclove, 1987). These information criteria statistics often show large initial decreases followed by more gradual decreases, even plateauing, with the addition of more (potentially uninformative) classes. Following the recommendations of Landa et al. (2012) and Petras and Masyn (2010), we considered the class size that yielded the final substantial decrease before the plateau to be the best fitting model, as determined by information criteria.

After determining the number of classes, we performed latent class regression to determine which baseline covariates were associated with latent class membership. We conducted this analysis in Mplus by fitting a joint model to the latent class indicators and covariates. Modal assignment, which assigns individuals to the latent class with the largest estimated probability of class membership, was used to estimate class-specific descriptive statistics, given the inherently unobservable nature of latent classes. All models were estimated using Mplus version 6.12 (Muthén & Muthén, 2010), which obtains parameter estimates using full information maximum likelihood (FIML) estimation, thereby adjusting for missing data. FIML methods assume data are missing at random, conditional on the variables in the model (Donders et al., 2006; Little, 1995; Muthén, 2004).

2.4 Results

Table 2.1 presents characteristics of adolescents included in this study. The mean age was 15.6 years, approximately 73% were male, 52% identified as White, 22% as Hispanic, 14% as Black and 12% as Other race. Youth were generally low to moderate substance users, reporting an average of 10.5 days of use in the prior 90 days; 28% of youth reported daily use. In the past year, adolescents exhibited an average of 2.4 DSM-IV symptoms of substance dependence. Nearly 75% of youth had no history of prior drug treatment; 51% of youth reported involvement with the juvenile or criminal justice systems (including incarceration, parole and probation) in the prior 90 days. Although all youth were community-residing at baseline, 33% reported having spent any time in a controlled environment (such as a jail, prison, or residential psychiatric facility) in the past 90 days.

Table 2.1 also reports the distribution of the 12 items that were used to define the latent classes. The majority of youth reported receiving individual-focused services, including relapse prevention training (78.1%), problem solving skills (90.7%), talking about friends (81.5%), and urine drug testing (70.4%). Fewer reported family-based or case management services. The most commonly reported family-based service was meeting at least twice with family members (58.5%), representing some family involvement in the treatment process. The most commonly reported case management service was calling the youth on the phone between appointments (56.8%), indicating some follow-up contact and scheduling reminders.

Table 2.2 and Figure 2.1 show the model fit statistics we considered for determining the optimal number of latent classes supported by the data. All information criteria statistics (AIC, BIC, and Adj BIC) only modestly decreased after 4 classes, indicating that this model is preferred (see "elbow" in Figure 1 plot; Landa et al., 2012; Petras and Masyn, 2010). The 3-class model had the highest entropy statistic (0.79) of all classes, indicating the least misclassification; the 4-class model had the next highest entropy statistic (0.78). Guided by the information criteria statistics, entropy and the relative interpretability of the 3- and 4-class models, we ultimately selected the 4-class model as the best fit.

Figure 2.2 shows the estimated proportions of youth in each class, as well as the probabilities of endorsing each item, given class membership. The latent classes are differentiated by varying combinations of treatment services, as well as differing probabilities of receiving given services. We describe the four classes as follows: Class 1: Low Service Utilization (12% of youth); Class 2: Individual-Focused Services (39%); Class 3: Individual- and Family-Focused Services (38%); and Class 4: Multiple Services (11%). In each class, individual-focused services were the services that youth were the most likely to receive; youth in Classes 1 and 2 (51% of the sample) were receiving primarily individual-focused services. The Low Service Utilization class had the lowest estimated item probability across all classes for 9 of the 12 services. The most common services for youth in this class were urine drug testing (57%) and problem solving training (38%); probabilities for all other services were less than 30%. Adolescents in the Individual-Focused Services class were likely to receive all four of the individual-focused services (item probabilities all >70%), and about half reported having treatment providers call them on the phone between appointments (51%). Youth in the Individual- and Family-Focused Services class likely received all four individual-focused services (item probabilities all >70%) as well as services that involved family members (96%) and worked on family communication (83%). In addition, youth in this class had a high likelihood of treatment providers calling them on the phone between appointments (64%). Finally, youth in the Multiple Services class were likely to receive the largest number of different services, with item probabilities for all twelve services greater than 65%. Several items had conditional probabilities exceeding 90% including instruction on relapse prevention, problem solving skill-building, discussing strategies for interacting with friends, family member involvement, and linking youth to other services.

Table 2.3 shows descriptive statistics by latent class, based on modal class assignment. In general, we observe a general gradation of substance use and other risk factors across classes that corresponds with the increasing number of treatment services defining Class 1 to Class 4. Across all classes, youth in the Low Service Utilization class report the lowest substance use, as measured by days of use (9.6 in past 90), the Substance Problems Scale, and the Substance Dependence Scale. Youth in this class were the least likely to have been in treatment before (22%) and exhibited the lowest motivation as assessed by the Treatment Motivation Index. Conversely, youth in the High Service Utilization class had the highest means on all indicators of substance use and were most likely to have been in treatment before (37%). These youth also showed the highest levels of mental health problems, the greatest involvement with the criminal justice system (67%), the lowest percentage of White adolescents (33%), the highest percentage of Hispanic adolescents (39%), and the highest treatment motivation. Additionally, the heterogeneity of services received in practice by youth in the same treatment program is apparent from the distribution of program enrollment across the 4 treatment classes. Each of the 4 classes contains youth from each of the 9 treatment programs.

Table 2.4 shows the pairwise odds ratios (i.e., the odds of being in one class relative to another) from our multivariate latent class regression and highlights many significant differences across classes. Youth with higher scores on the Substance Problem Scale were significantly less likely to be in the Low Services class relative to each of the other classes (Low v: Indiv OR= 0.90; Indiv & Fam OR=0.85; Mult OR=0.89), as were youth with higher treatment motivation (Low v: Indiv OR=0.85; Indiv & Fam OR= 0.83; Mult OR=0.76). Youth with higher scores on the Substance Dependence Scale were more likely to be in the Low Services class compared to the Individual-Focused class (OR = 1.18) and the Individual- and Family-Focused class (OR=1.25).

Youth in the Individual-Focused Services class were significantly more likely to be older (Indiv v: Low OR=1.26; Indiv & Fam OR=1.19; Mult OR=1.15); less likely to have spent time in a controlled environment (Indiv v: Low OR=0.75; Indiv & Fam OR=0.63; Mult OR=0.62); and less likely to have been arrested (Indiv v: Indiv & Fam OR=0.70; Mult OR=0.71). Additionally, they were less likely to recognize their substance problems (Indiv v: Low OR=0.62; Indiv & Fam OR=0.53; Mult OR=0.65); less likely to have been in treatment before (Indiv v: Indiv & Fam OR=0.81; Mult OR=0.69); and less likely to be Black, compared to White (Indiv v: Low OR=0.50; Mult OR=0.40).

Youth in the Individual- and Family-Focused Services class were more likely to be Hispanic, relative to White, compared to youth in the Low Services and the Individual-Focused classes (Indiv & Fam v: Low OR=2.02; Indiv OR=1.90), yet less likely compared to youth in the Multiple Services class (Indiv & Fam v Mult OR=0.43). Youth with higher scores on the Substance Problems Scale had higher odds of being in this class relative to the Low Services and Individual-Focused classes (Indiv & Fam v: Low OR=1.18; Indiv OR=1.07).

Minority youth had elevated odds, relative to White youth, of being in the Multiple Services class compared to other classes. Specifically, Hispanic youth were more likely to be in this class compared to all other classes (Mult v: Low OR=4.70; Indiv OR=4.41; Indiv & Fam OR=2.33), Black youth were more likely relative to the Individual-Focused and the Individual- and Family-Focused Services classes (Mult v: Indiv OR=2.52; Indiv & Fam OR=1.92), and Other race youth were more likely relative to be in the Low Services and the Individual- and Family-Focused Services classes (Mult v: Low OR=1.66; Indiv & Fam OR=1.54). Youth with juvenile/criminal justice system involvement had higher odds of being in the Multiple Services class (Mult v: Low OR=1.43; Indiv OR=1.62; Indiv & Fam OR=1.59), as did youth with higher treatment motivation (Mult v: Low OR=1.31; Indiv & Fam OR=1.12). Overall, our latent class regression highlights that there are significant differences in youth receiving different types of treatment services.

2.5 Discussion

This study demonstrates that, among a large sample of adolescents receiving outpatient services through various treatment programs in the US, distinct groups of youth can be empirically identified based on the services they received. Four classes were identified: Low Service Utilization (12% of youth), Individual-Focused Services (39%), Individual- and Family-Focused Services (38%), Multiple Services (11%). The combinations of services that define these classes are recognizable and highly plausible given that individual-focused services, such as those provided through CBT therapy, are the most likely services within each class; more integrated, comprehensive treatment approaches expand upon individual-focused services by adding family-based or case-management components.

Importantly, program enrollment did not accurately reflect services ultimately received by youth in this data. Each of the 9 treatment programs had youth estimated to be in all 4 latent classes. Differences in service provision within a given treatment program may be due to differences in youth need, noncompliance, dropout, service wait-listing, or may reflect services received through additional treatment providers. Since program enrollment and service provision were distinct in this data, the typical approach of classifying youth based on program enrollment would fail to accurately delineate youth with respect to services actually received. Furthermore, the fact that 9 treatment programs can be well represented by only 4 classes highlights commonalities in service provision across programs. Thus, a latent class approach represents a parsimonious approach to identifying treatment groups that may prove advantageous when estimating treatment effects in data representing individuals in numerous treatment programs.

It is important to highlight that our class definitions only provide information as to the likelihood that a youth *ever* received a given service, rather than service frequency. Receiving a smaller number of services types should not be conflated with receiving less treatment overall, nor should treatment classes be interpreted as representing a hierarchy. Also, the low item probabilities in the Low Services class imply that some youth in this class may not have received any services at all, despite being enrolled in a treatment program. A number of factors, including treatment refusal, dropout, treatment wait-listing, or inability to continue treatment due to incarceration, may have resulted in youth enrolled in treatment programs receiving no or very few services.

As expected given the observational nature of the data, significant differences in youth characteristics were observed across classes. For example, youth in the Low Service Utilization class exhibited lower scores on the Substance Problems Scale, and may have received fewer service types in part due to lower treatment need. Additionally, youth in the Multiple Services class were the most likely to have been involved in the justice system or to have spent time in a controlled environment. This association reflects the fact that some youth were receiving services through the justice system or a community re-entry program; three of the 9 treatment programs included in this study, the Strengthening Communities' Youth program, the Juvenile Treatment Drug Court program, and the Young Offenders Reentry Program, were specifically designed for youth with justice system involvement.

Race/ethnicity was also associated with treatment class, such that minority youth had elevated odds of being in the Multiple Services class relative to each of the other classes. The cause of these demographics trends is not entirely clear from our data. Race/ethnicity was also associated with justice system involvement, and since youth involved in justice systems also have higher odds of being in the Multiple Services class, these demographic trends may represent residual associations of race/ethnicity and justice system involvement. Alternatively, these trends may be a reflection of varying degrees of access to treatment services across racial/ethnic groups.

This study utilizes a large, multisite dataset of adolescents receiving substance use treatment in the US and introduces a novel approach, using latent class analysis, of empirically identifying common clusters of treatment services youth receive across numerous treatment programs and sites. Although the GAIN survey provides a rich set of well-validated variables relating to substance use, treatment, and risk factors, one limitation is the self-reported nature of the data. The accuracy of self-reported data may be affected by recall bias, social desirability bias, or other factors. Unfortunately, we do not have alternative sources, such as provider reports or administrative data, to verify youth reports; it is possible that youth may be either over- or under-reporting services, thereby leading to misclassification of some youth with respect to treatment classes. Additionally, we lack information regarding the frequency of services received, quality metrics of services, and the authority providing the services; such data could enrich the descriptions of our latent classes. Finally, the descriptions of treatment classes in our study may not be fully generalizable to other populations of adolescents receiving outpatient treatment services.

Two important clinical and policy objectives for future research that arise from this study include determining whether services received in typical outpatient drug treatment are well-matched to youths' needs, as well as examining the relative effectiveness of these classes of treatment services on substance use. Many factors determine which treatment services would be an ideal fit for a given adolescent, including severity and nature of substance use, family environment, motivation for change, personality, and educational status. Although we cannot assess the match between need and services in our data, there is evidence from the adult drug treatment literature that outcomes are improved when services match client need and preference. Furthermore, it is of significant practical importance to conduct evaluation studies examining the effectiveness of services that youth are actually receiving, rather than the services that treatment programs purport to provide. As demonstrated by our analysis, in such an analysis it will be imperative to appropriately account for heterogeneity in factors such as demographics, baseline substance use and justice system involvement across treatment classes. Having an accurate description of the services that adolescents are actually receive in the course of standard outpatient substance use treatment is the essential first step to investigating both of these important questions.

Table 2.1. Descriptive statistics of the total sample

	Overall Sample N=5527
	Mean or %
Demographics	
Age	15.61
Female	26.5%
White	51.6%
Black	13.7%
Hispanic	22.0%
Other	12.6%
Substance Use	
Daily Substance Use	28.2%
History of substance use treatment	26.8%
Days of substance use, past 90 days	10.46
Substance Problems Scale, past year	6.45
Substance Dependence Scale, past year	2.38
Treatment Motivation Index (Max=5)	1.83
Does not recognize AOD problems	11.7%
Legal (past 90 days)	
Criminal justice system involvement	51.3%
Spent time in controlled environment	32.9%
Arrested	21.4%
Crime Violence Scale (Max=31)	6.52
Days involved with illegal activities	9.55
Mental Health (past 90 days)	
Days affected by emotional problems	19.79
Internal Mental Distress Scale (Max=43)	7.42
Suicidal thoughts	10.6%
Behavior Complexity Scale (Max=31)	10.12
Problem Orientation Scale (Max=5)	0.68
Sought mental health treatment	17.6%
Environmental	
Living Environment Risk Scale (Max=28)	10.39
Social Environment Risk Index (Max=28)	13.2
Any school attendance, past 90 days	91.0%
Any employment, past 90 days	33.5%
Latent Class Indicators	
Individual-focused services	
Teach or review with you relapse prevention problems?	78.1%
Talk about different ways to solve problems?	90.8%
Talk with you about your friends?	81.5%
Require you to take urine tests?	70.4%
Family-based services	
Work with you at your home?	24.5%
Meet with family members of yours more than one time?	58.5%
Work with members of your family on communication?	46.0%
Hook your family up with services?	13.0%

Case management services	
Call you on the phone in between appointments?	56.8%
Talk with a counselor, teacher or other adult at school?	19.2%
Hook you up with other services?	33.3%
Provide you with transportation to appointments?	26.3%
Treatment Program Enrollment	
AAFT	9.8%
ART	0.9%
ATM	3.8%
CYT	8.2%
DC	6.2%
EAT	45.1%
SCY	14.3%
TCE	5.1%
YORP	6.5%

Abbreviations: AAFT=Assertive Adolescent and Family Treatment; ART=Adolescent Residential Treatment; ATM = Adolescent Treatment Models; CYT= Cannabis Youth Treatment; DC= Family and Juvenile Treatment Drug Court; EAT=Effective Adolescent Treatment; SCY= Strengthening Communities' Youth; TCE=Targeted Capacity Expansion; YORP= Young Offenders Reentry Program

Table 2.2. Fit statistics for latent class models with one to six classes

# Classes	*# Free Parameters*	*AIC*	*BIC*	*Adj BIC*	*Entropy*	*LMR LRT p-val*	*BLRT p-val*
1 Class	12	72871.92	72951.32	72913.19	--	--	--
2 Class	25	67420.31	67585.75	67506.30	0.75	0.0000	0.0000
3 Class	38	66000.45	66251.91	66131.16	0.79	0.0000	0.0000
4 Class	51	64855.84	65193.33	65031.26	0.78	0.0000	0.0000
5 Class	64	64588.58	65012.09	64808.72	0.77	0.0002	0.0000
6 Class	77	64318.32	64827.86	64583.18	0.76	0.0000	0.0000

Abbreviations: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; Adj BIC = sample size adjusted Bayesian Information Criterion; LMR LRT = Lo-Mendel-Rubin likelihood ratio test; BLRT = Bootstrap likelihood ratio test.

Figure 2.1. Plot of information criteria fit statistics for 1 – 6 class models

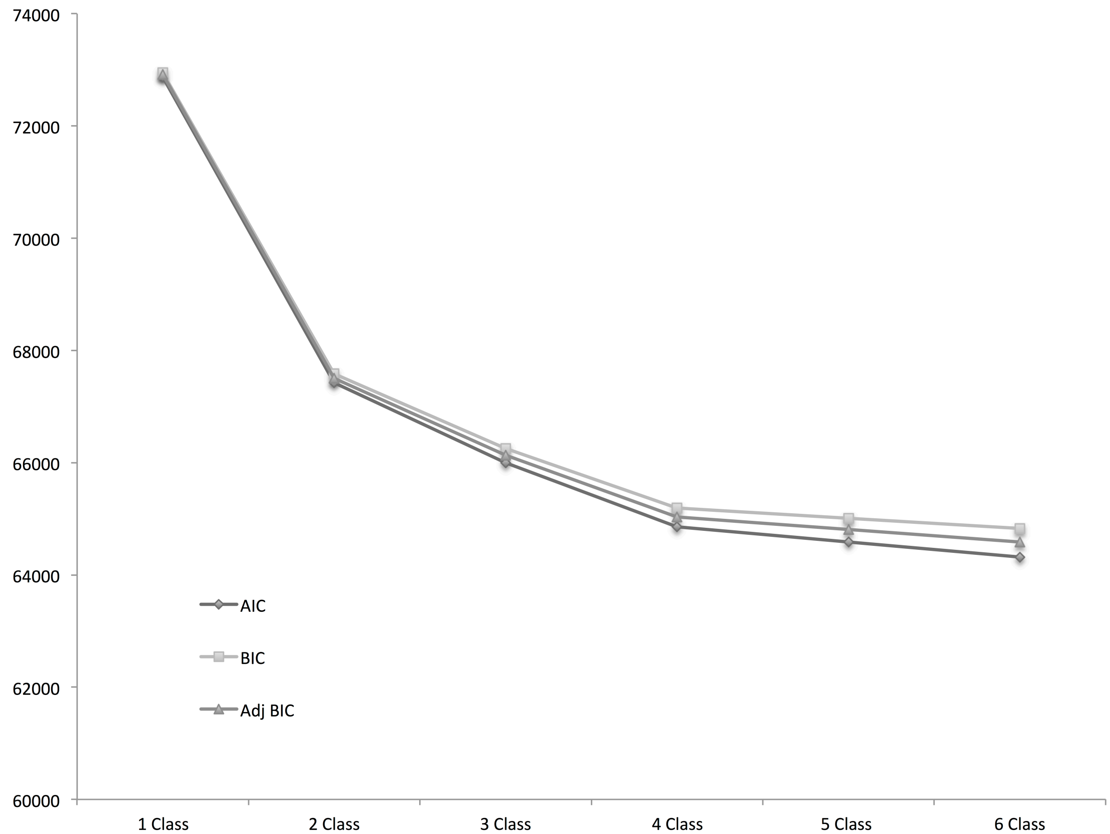


Figure 2.2. Conditional item probabilities based on a 4-class latent class analysis (n=5527)

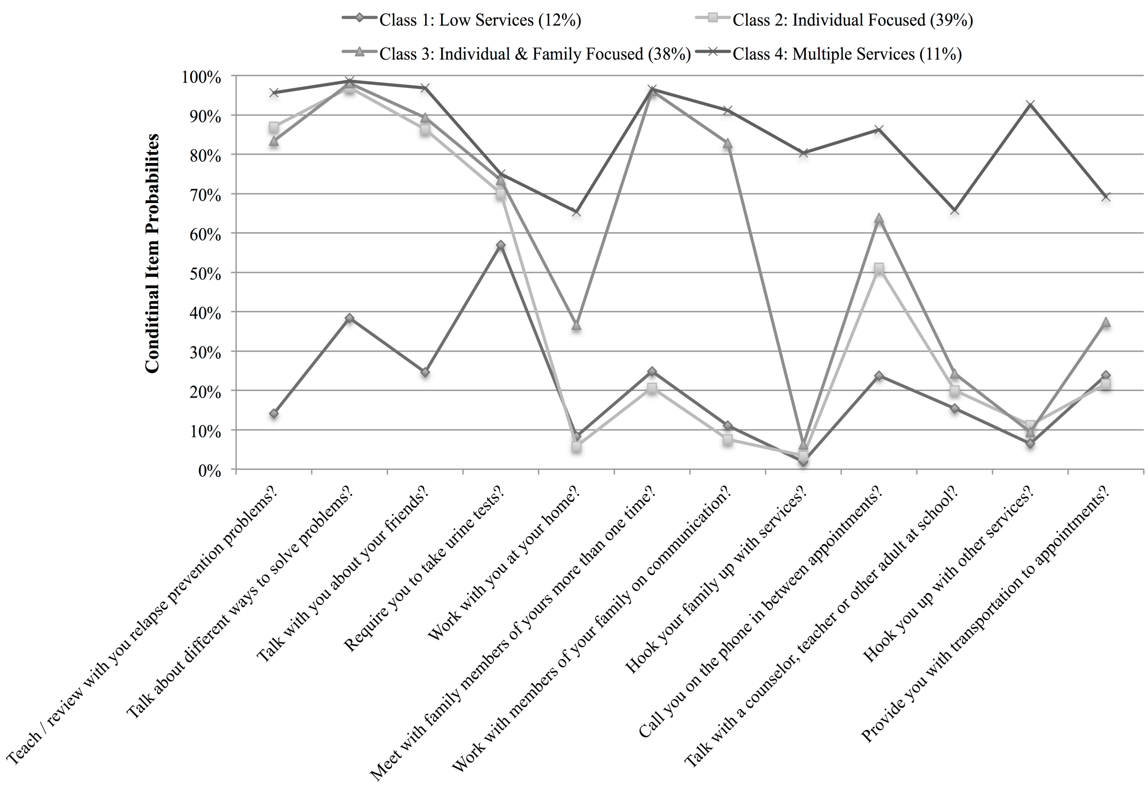


Table 2.3. Descriptive statistics by latent class

	Class 1: Low Service Utilization N=579	*Class 2: Individual Focused N=2340*	*Class 3: Individual & Family Focused N=2016*	*Class 4: Multiple Services N=592*
	Mean or %	**Mean or %**	**Mean or %**	**Mean or %**
Demographics				
Age	15.48	15.71	15.54	15.56
Female	26.3%	28.5%	24.3%	26.4%
White	52.0%	57.8%	49.8%	33.3%
Black	21.8%	11.9%	12.7%	16.4%
Hispanic	15.4%	15.6%	26.3%	39.4%
Other	10.9%	14.7%	11.2%	11.0%
Substance Use				
Daily Substance Use	26.3%	25.3%	30.4%	33.4%
History of substance use treatment	21.8%	23.5%	29.0%	36.9%
Days of substance use, past 90 days	9.6	9.66	11.35	11.43
Substance Problems Scale, past year	5.5	6.14	6.88	7.2
Substance Dependence Scale, past year	1.99	2.24	2.55	2.77
Treatment Motivation Index (Max=5)	1.57	1.74	1.93	2.06
Does not recognize AOD problems	13.3%	9.4%	14.1%	11.0%
Legal (past 90 days)				
Criminal justice system involvement	50.9%	45.6%	53.5%	66.9%
Spent time in controlled environment	31.6%	25.6%	37.3%	48.1%
Arrested	20.0%	18.0%	24.6%	25.3%
Crime Violence Scale (Max=31)	6.18	6.3	6.64	7.34
Days involved with illegal activities	8.44	8.94	10.33	10.36
Mental Health (past 90 days)				
Days affected by emotional problems	19.56	19.83	19.72	20.05
Internal Mental Distress Scale (Max=43)	6.9	7.17	7.56	8.46
Suicidal thoughts	10.2%	10.7%	10.8%	9.8%
Behavior Complexity Scale (Max=31)	9.34	9.67	10.63	10.93
Problem Orientation Scale (Max=5)	0.45	0.64	0.73	0.9
Sought mental health treatment	15.2%	18.7%	17.0%	18.1%
Environmental				
Living Environment Risk Scale (Max=28)	10.4	10.28	10.48	10.49
Social Environment Risk Index (Max=28)	13.09	12.96	13.4	13.6
Any school attendance, past 90 days	91.3%	92.0%	90.8%	87.9%
Any employment, past 90 days	31.4%	37.2%	32.4%	24.6%
Latent Class Indicators				
Individual-focused services				
Teach or review with you relapse prevention problems?	14.1%	87.0%	83.4%	95.6%
Talk about different ways to solve problems?	38.4%	97.0%	98.0%	98.6%
Talk with you about your friends?	24.6%	86.4%	89.3%	96.8%
Require you to take urine tests?	57.0%	70.1%	73.5%	75.0%
Family-based services				
Work with you at your home?	8.3%	5.8%	36.6%	65.4%
Meet with family members of yours more than one time?	24.8%	20.7%	96.1%	96.5%
Work with members of your family on	11.1%	7.6%	82.9%	91.2%

communication?				
Hook your family up with services?	1.9%	3.4%	6.3%	80.4%
Case management services				
Call you on the phone in between appointments?	23.8%	51.2%	63.8%	86.2%
Talk with a counselor, teacher or other adult at school?	15.5%	20.0%	24.3%	65.8%
Hook you up with other services?	6.5%	11.2%	9.4%	92.6%
Provide you with transportation to appointments?	23.9%	21.7%	37.4%	69.2%
Treatment Program Enrollment				
AAFT	6.0%	5.2%	13.8%	18.2%
ART	0.9%	0.4%	1.1%	2.4%
ATM	4.0%	2.6%	5.1%	4.2%
CYT	10.0%	5.3%	11.4%	6.8%
DC	1.9%	2.0%	9.7%	15.7%
EAT	41.1%	61.4%	31.3%	31.9%
SCY	20.4%	12.4%	15.2%	12.8%
TCE	6.4%	5.4%	5.3%	2.0%
YORP	9.3%	5.3%	7.1%	5.9%

Abbreviations: AAFT=Assertive Adolescent and Family Treatment; ART=Adolescent Residential Treatment; ATM = Adolescent Treatment Models; CYT= Cannabis Youth Treatment; DC= Family and Juvenile Treatment Drug Court; EAT=Effective Adolescent Treatment; SCY= Strengthening Communities' Youth; TCE=Targeted Capacity Expansion; YORP= Young Offenders Reentry Program
Note: Class-specific statistics were estimated using modal class assignment

Table 2.4. Pairwise odds ratios (OR) of class membership from multinomial latent class regression

	Indiv v Low Services		***Indiv + Fam v Low Services***		***Multiple v Low Services***	
	OR	**p-value**	**OR**	**p-value**	**OR**	**p-value**
Female	1.16	0.27	0.85	0.19	1.11	0.53
Age	1.26	0.00*	1.05	0.23	1.10	0.08
Black	0.50	0.00*	0.66	0.01*	1.26	0.25
Hispanic	1.07	0.72	2.02	0.00*	4.70	0.00*
Other	1.40	0.06	1.08	0.65	1.66	0.02*
History of substance use treatment	0.90	0.47	1.11	0.46	1.30	0.10
Substance Problems Scale, past year	1.10	0.01*	1.18	0.00*	1.12	0.01*
Substance Dependence Scale, past year	0.85	0.03*	0.80	0.00*	0.89	0.17
Treatment Motivation Index	1.17	0.00*	1.21	0.00*	1.31	0.00*
Does not recognize AOD problems	0.62	0.01*	1.17	0.32	0.96	0.84
Criminal justice system involvement	0.89	0.33	0.90	0.40	1.43	0.03*
Spent time in controlled environment	0.75	0.05*	1.20	0.17	1.44	0.03*
Arrested	0.83	0.23	1.19	0.21	1.16	0.39
Crime Violence Scale	1.01	0.58	0.98	0.07	0.99	0.60
Days affected by emotional problems	1.00	0.90	0.99	0.04*	0.99	0.22
Behavior Complexity Scale	0.99	0.32	1.02	0.09	1.01	0.70
	Indiv + Fam v Indiv		***Multiple v Indiv***		***High Services v Indiv + Fam***	
	OR	**p-value**	**OR**	**p-value**	**OR**	**p-value**
Female	0.73	0.00*	0.96	0.72	1.31	0.04*
Age	0.84	0.00*	0.87	0.00*	1.05	0.32
Black	1.31	0.05	2.52	0.00*	1.92	0.00*
Hispanic	1.90	0.00*	4.41	0.00*	2.33	0.00*
Other	0.78	0.04*	1.19	0.34	1.54	0.03*
History of substance use treatment	1.23	0.03*	1.45	0.00*	1.18	0.20
Substance Problems Scale, past year	1.07	0.01*	1.02	0.66	0.95	0.16
Substance Dependence Scale, past year	0.94	0.19	1.04	0.57	1.11	0.16
Treatment Motivation Index	1.03	0.35	1.12	0.02*	1.09	0.08
Does not recognize AOD problems	1.88	0.00*	1.54	0.02*	0.82	0.27
Criminal justice system involvement	1.02	0.82	1.62	0.00*	1.59	0.00*
Spent time in controlled environment	1.59	0.00*	1.91	0.00*	1.21	0.17
Arrested	1.43	0.00*	1.40	0.02*	0.97	0.83
Crime Violence Scale	0.97	0.00*	0.98	0.21	1.02	0.21
Days affected by emotional problems	0.99	0.01*	0.99	0.14	1.00	0.78
Behavior Complexity Scale	1.03	0.00*	1.02	0.14	0.99	0.23

* denotes p-value < 0.05

CHAPTER 3. One-step and three-step methods for categorical latent variable regression

3.1 Abstract

Background: Latent class analysis is a common statistical method used by social and behavioral researchers; latent variable regression is an extension that estimates the association between latent class and an auxiliary variable. Two common types of models for latent variable regression are those that (1) regress a latent class outcome on an observed predictor (e.g., treatment); and (2) regress an observed distal outcome on latent treatment classes. Broadly, these models can be estimated using 1-step methods, classical 3-step methods, or more recent corrected 3-step methods; this paper reviews the conceptual differences between methods and compares their statistical performance using simulated data.

Methods: We perform a simulation study in which we compare a 1-step method, three classical 3-step methods and a corrected 3-step method for latent variable regression with a distal outcome. Data with a 3-class structure, 15 binary class indicators, and a continuous outcome variable were generated in Mplus. We considered conditions with varying levels of entropy, class separation in terms of outcomes (contrast size), and variance in the distal outcome. One-step estimation was performed in R using the LCCA package; modal assignment, multiple pseudoclass assignment, and posterior probability regression were also implemented in R. The corrected 3-step method was implemented in Mplus v7.11. The estimates of interest were the pairwise differences in outcome means across classes; statistical performance was assessed in terms of bias, standard error, mean squared error, and 95% confidence interval coverage.

Results: The 1-step method and the corrected 3-step method both performed quite well across all conditions we investigated. Modal and pseudoclass assignment yielded significantly biased estimates, attenuated standard errors, and poor coverage rates, particularly when entropy was low.

Conclusions: One-step methods perform well with respect to latent variable regression; however, these methods may not always converge due to model complexity or may not always be conceptually appropriate. Although commonly used by applied researchers, classical 3-step methods often perform poorly in practice. Thus, when a 3-step approach is desired, recent corrected 3-step methods should be used.

3.2 Introduction

Latent variable modeling is now a very common statistical approach in social science, public health, and prevention research, given that these fields all frequently seek to quantify constructs (e.g., socioeconomic status, substance use phenotypes, risky sexual behavior) that are not fully observable. Since these constructs cannot be assessed through any one direct item, often instruments with many items are used, with the fundamental assumption that the observed scale reflects the underlying latent construct. However, appropriate statistical methods are required to account for the measurement error that is associated with the unobservable nature of the latent variable. Additionally, there is often interest in estimating the association between a latent variable and auxiliary variables, either predictors or outcomes. Of particular interest are models that (1) treat the latent variable as the outcome and model the association between predictors and latent class membership; or (2) treat the latent variable as the predictor and model the association between latent class membership and outcomes. An example of the former is a recent study examining the link between early conduct problems and latent classes of risky sexual behavior during adolescence (Conduct Problems Prevention Research Group, 2013). An example of the later is a study by Feingold et al. (2013) investigating the association of latent classes of substance use with subsequent acts of intimate partner violence.

Despite their widespread use, there are many unresolved methodological challenges regarding how best to estimate latent variable regression models; the dominant estimation methods can be classified as either 1-step or 3-step methods. One-step methods, which jointly estimate the latent variable measurement model and the regression structural model, are optimal with regard to bias and efficiency, yet a 1-step model may not always be feasible due to computational complexity. Three-step methods, which separately estimate the latent variable measurement model and the regression structural model, are commonly used but have been shown to yield biased estimates for the association between latent variables and auxiliary variables. More recently, methods that correct the bias inherent to classical 3-step methods have been proposed; these methods perform similarly to 1-step methods but may be easier to implement under some conditions. However, new corrected 3-step methods have not yet been widely disseminated among or adopted by applied researchers.

Previous studies have compared 1-step and 3-step methods using simulated data (e.g., Asparouhov & Muthén, 2013; Bolck et al., 2004; Clark & Muthén, 2009; Petersen et al., 2012; Vermunt, 2010); however, these studies have primarily been written at a more technical level and not intended for dissemination among applied researchers. Recently, Feingold et al. (2013) gave a nice overview this issue for applied researchers, yet compared methods using actual data, making it difficult to objectively compare performance of methods since true values for estimated parameters are not known. The objective of this paper is to provide a comprehensive overview of 1-step, classical 3-step, and corrected 3-step methods for latent variable regression for applied researchers, emphasizing that 1-step or corrected 3-step methods should be used. We compare 5 different methods using simulated data and compare a broader range of simulation conditions than previous studies.

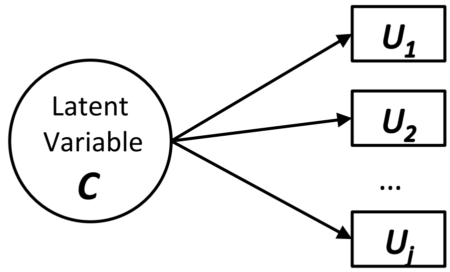
3.3 Latent class analysis

In general, latent variable modeling is appropriate when there is reason to believe that the population of interest has an underlying structure defined by a latent construct that is not directly observable but rather indirectly measured by a collection related indicator items that pertain to different aspects of underlying latent construct. Latent class analysis (LCA) is a type of finite mixture modeling that is used to identify discrete and mutually exclusive subgroups of individuals within a population, based on observed response patterns on a set of indicator items (Collins & Lanza, 2010; Goodman, 1974; Haberman, 1979; Lazarsfeld & Henry, 1968; McCutcheon, 1987). Latent class analysis assumes that this structure is defined by mutually exclusive and homogenous latent classes; the underlying latent variable C is modeled as categorical, such that each individual belongs to exactly one of C latent classes. Latent class indicators, denoted $U_1, U_2, \ldots, U_j$, can be binary or categorical indicators.

This paper focuses on latent class analysis, which models a categorical latent variable from cross-sectional categorical indicators. There are several other types of categorical latent variable modeling, including latent profile analysis (LPA), which models a categorical latent variable from cross-sectional continuous latent variables, and growth mixture modeling (GMM), which models a categorical latent variable from longitudinal, typically continuous, indicators. Although not discussed in this paper, corrected 3-step methods similar to those discussed in this paper have been introduced for LPA (Gudicha & Vermunt, *in press*) and GMM (Asparouhov & Muthén, 2013; Muthén & Muthén, 2012). Models for continuous latent variables, such as factor analysis (cross-sectional continuous indicators) and latent trait analysis (cross-sectional categorical indicators), also are subject to bias under 3-step methods. Corrected methods for continuous latent variable regression are discussed in Croon (2002), Lu & Roland (2008), Oberski & Satorra (2013), and Skrondal & Laake (2001).

In the parlance of structural equation modeling, a latent class analysis model is strictly a *measurement* model since it lacks a *structural* model component. A measurement model describes the relationship between the observed indicator items and the latent variable. The structural model describes the relationship between the latent variable and auxiliary variables (either observed or latent). Figure 3.1 represents the model for latent class analysis.

Figure 3.1. Latent class analysis model. $U_1, U_2, \ldots, U_j$ represent latent class indicators manifested by the latent variable, denoted C.



When conducting LCA, two types of parameters are of interest: conditional item probabilities (denoted as ρ) and posterior class membership probabilities (denoted as γ). The conditional item probability represents the probability of endorsing response r_j for item U_j, conditional on membership in class c, and is defined as $\rho_{j,r=rj|c} = \Pr\left(U_j = r_j \mid C = c\right)$. The posterior class membership probability represents the probability that an individual belongs to latent class c, given his/her observed response pattern $\boldsymbol{u}= (U_1=r_1, U_2=r_2,\ldots,U_j=r_j)$, and is denoted as $\gamma_{c|u} = \Pr(C = c|\boldsymbol{U} = \boldsymbol{u})$. Posterior probabilities can be used to estimate the class prevalences, namely the proportion of the total sample predicted to belong to each of the C latent classes, denoted $\pi_c = \Pr(C = c)$ for $c = 1,\ldots,C$. The classical latent class analysis model can be expressed as follows: $\Pr(\boldsymbol{U} = \boldsymbol{u}) = \sum_{\boldsymbol{c=1}}^{\boldsymbol{C}} \Pr\,(C = c) \prod_{j=1}^{J} \Pr\left(U_j = r_j \mid C = c\right)$.

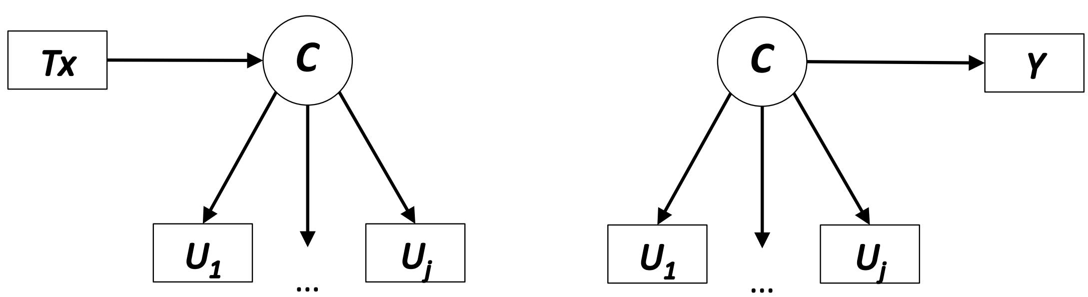
A central assumption of latent class analysis is local independence of the indicators; that is, the correlation between indicators is entirely explained by the latent class, such that there are no residual correlations among the indicators after conditioning on latent class membership (Lazarsfeld & Henry, 1968). This assumption is represented in Figure 3.1 by the absence of any correlational arrows between the Us. Additionally, latent class analysis assumes that all individuals are independent, such that one individual's class membership does not affect others' class memberships.

3.4 Latent structure models

Generally, public health and prevention researchers are interested in investigating the relationships among a latent variable and auxiliary variables of interest, a class of models known as latent structure models (Lazarsfeld & Henry, 1968; Clogg, 1995; Croon, 2002). These auxiliary variables may either be observed or latent variables and may either be predictors of the latent variable (i.e., independent variables) or predicted by the latent variable (i.e., dependent variables). In this article we focus only on methods that relate categorical latent variables to observed auxiliary variables, specifically, regressing a latent variable on covariates or a distal outcome on a latent variable (see Figure 3.2).

These models require all of the assumptions of latent class analysis, as well as the additional assumption of conditional independence, which assumes the latent variable indicators are independent of the auxiliary variable. This assumption is denoted in Figure 3.2 by a lack of a direct effect arrow connecting the indicators *U*s and *X* in Panel A by lack of a direct effect arrow connecting the indicators *U*s and *Y* in Panel B.

Figure 3.2. Schematic representation of latent variable regression models considered in this paper.



Panel A. Latent class regression with predictive treatment

Panel B. Latent class regression with distal outcome

Broadly, latent class regression models can be estimated with either a 1-step or 3-step approach. A 1-step approach fits a joint model and estimates the measurement and structural models simultaneously – described as a "summarize AND analyze" approach by Clogg (1995). In contrast, a 3-step approach first estimates the measurement model, uses the estimated posterior probabilities to construct observed variables representing class membership, and finally estimates the structural part of the model – a "summarize THEN analyze" strategy.

3.5 One-Step models

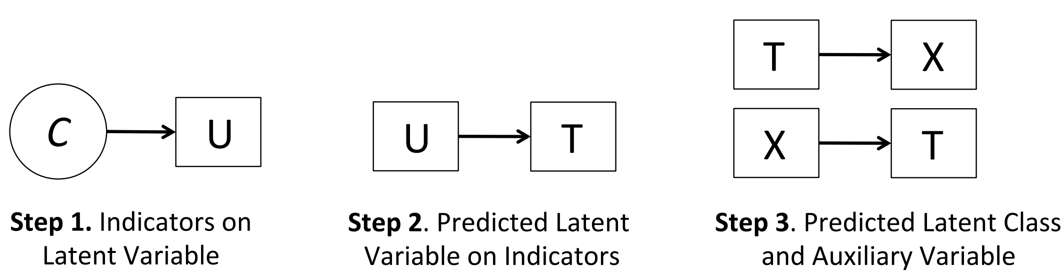
By fitting a joint model that incorporates both the latent class measurement model and the structural model, the 1-step approach allows simultaneous estimation of the parameters of interest. One-step estimation uses full information likelihood estimation, and thus yields unbiased, efficient and consistent parameter estimates, assuming the model is correctly specified. Maximum likelihood estimates, based on the joint likelihood, are typically estimated using the E-M algorithm (Bandeen-Roche et al., 1997; Dayton & Macready, 1988; Goodman, 1974). In the joint model, parameters in the structural part of the model are estimated based on the posterior probabilities of class membership, thereby appropriately accounting for the uncertainty in class membership when calculating parameter standard errors. Additionally, joint modeling allows the auxiliary variables (i.e., predictive covariates or distal outcomes) to influence the estimation of the measurement model, which can lead to improved efficiency. For example, for latent class regression, joint modeling can be desirable when there is relatively poor class separation, since including covariates in the model can better help to distinguish classes.

A notable drawback to this approach is the potential for computational complexity – not all statistical packages have the capability to implement such models, and complex models may suffer from convergence issues. Also, misspecification of the measurement part of the model may bias parameter estimates for the structural part, and misspecification of the structural part may bias measurement parameters (Bolck et al., 2004). Another drawback is that in a joint model, the interpretation of the classes must involve the auxiliary variables as well as the indicators; opinions differ on the interpretability of such classes. Additionally, there are concerns from a causal inference perspective on fitting a joint model in the case of distal outcomes (Petras & Maysn, 2010). This allows the distal outcome, measured subsequently to the indicators, to influence the latent class structure, thereby ignoring the temporal ordering of the variables. If the goal is to estimate the predictive ability or causal effect of the latent classes, the temporal ordering must be preserved; thus, a 3-step model, which would estimate latent classes independent of the distal outcome, may be preferable.

3.6 Classical 3-Step Methods

The first step in a classical 3-step method is to estimate the measurement model using a latent class analysis (without covariates). The second step uses the results from the LCA to generate an observed variable representing class membership for each individual, discussed in detail below. One approach is to directly use the estimated posterior probabilities, $\gamma_{c|u}$ for c =1, …, C. An alternative approach is to use the posterior probabilities to generate a categorical variable representing predicted latent class (also called the latent score) for each individual. The third step estimates the association between the auxiliary variable of interest and latent class, either described by posterior probabilities or predicted latent class.

Figure 3.4. Schematic representation of steps in a 3-step analysis for latent variable regression.



Unlike the 1-step method, the classical 3-step method does not directly estimate the association between the auxiliary variable and the true latent variable C; rather, it estimates the association between the auxiliary variable and an observed proxy variable for C. As Vermunt

(2010) highlights, most applied researchers do not think in terms of joint models, and thus a 3-step approach reflects the practice of first building the measurement model and then investigating the relationships between latent classes and auxiliary variables. Additionally, 3-step methods may offer estimation advantages in the case of data sparseness. One-step methods may suffer when particular combinations of latent classes and indicators have small cell sizes (sample sizes) with respect to the auxiliary variable. Fitting a 3-step model may circumvent this problem somewhat, since fitting the structural model relies on the joint distribution of the predicted latent scores and the auxiliary variable, not the joint distribution of the latent scores, latent variable indicators and the auxiliary variable. See Bolck et al. (2004) and Bakk et al. (2013) for more discussion.

Modal / Maximum-probability assignment: One approach for the 2nd step in a 3-step approach is to use modal assignment, also called maximum-probability assignment (Clogg, 1995). This is a deterministic approach under which each individual with response pattern $\boldsymbol{u}$ is assigned to the class in which they have the highest posterior probability of being in (i.e., the class for which $\gamma_{c|u}$ is maximal across all c). With modal assignment, the structural model in Step 3 estimates the association between modal class and the auxiliary variable, using standard regression methods. If modal class is a predictor of a distal outcome, it is treated as a polytomous predictor variable; if modal class is the outcome, multinomial regression is used.

Pseudoclass (random) assignment: Random assignment, also called pseudoclass assignment, is a probabilistic approach under which each individual with response pattern $\boldsymbol{u}$ is assigned a predicted class based on a random draw from a multinomial distribution specified by the posterior probabilities $\gamma_{u,1}$, $\gamma_{u,2},\ldots, \gamma_{u,C}$. This approach may either use a single draw or multiple (e.g., 20) draws (Bandeen-Roche et al., 1997; Goodman, 2007; Wang et al., 2005). Similar to modal assignment, the structural model in Step 3 estimates the association between pseudoclass and the auxiliary variable using standard regression methods. As for modal assignment, if pseudoclass is a predictor of a distal outcome, it is treated as a polytomous predictor; if pseudoclass is the outcome, multinomial regression is used. When using a multiple pseudoclass approach with k draws, k separate structural models are fit and the combining rules for multiple imputation are used to generate the final parameter estimates (Rubin, 1987).

Probability regression: This approach uses the posterior probabilities $\gamma_{u,1}$, $\gamma_{u,2}$,…, $\gamma_{u,C}$ directly to represent latent class membership. When the auxiliary variable is a predictor, the structural model in Step 3 regresses a vector of $C - 1$ posterior probabilities on the auxiliary variable. A log linear model is recommended in order to obtain properly bounded posterior probability estimates (Clark & Muthén, 2009). Petersen et al. (2012) give details of one estimation approach using a non-linear multivariate normal model with an unstructured covariance matrix. When the auxiliary variable of interest is a distal outcome, the structural model in Step 3 regresses the auxiliary variable on $C - 1$ posterior probabilities.

3.7 Limitations of classical 3-step methods

When estimating the association between an auxiliary variable and latent class, there are two significant limitations – misclassifying individuals with respect to true latent class by using a proxy variable, and improperly accounting for the uncertainty of class membership.

Misclassification of individuals with respect to their true latent class is particularly a problem for modal and pseudoclass assignment; posterior probability regression avoids much of this bias by actually using the posterior probabilities themselves. Goodman (2007) shows that the misclassification under pseudoclass assignment will always be greater than under modal assignment, which by design minimizes the number of individuals who are misclassified. Recall that the latent class model estimates a set of posterior probabilities ($\gamma_{1|u}$, $\gamma_{2|u}$,…, $\gamma_{C|u}$) for each response pattern $\boldsymbol{u}$. Under modal assignment, all n_u individuals with a given response pattern $\boldsymbol{u}$ are predicted to be in the class that corresponds to the highest posterior probability for that response pattern, denoted by class m. On average, ($\gamma_{m|u}$)% of these individuals truly belong to the most likely class, and the rest the are misclassified, yielding a predicted misclassification rate of

$(1 - \gamma_{m|u})$% for this response pattern. Under pseudoclass assignment, individuals with a given response pattern $\boldsymbol{u}$ are each assigned k predicted classes via random draw from the multinomial probability distribution defined by $(\gamma_{1|u}, \gamma_{2|c}, \ldots, \gamma_{C|u})$. Of the n_u individuals with this response pattern, on average, $(n_u \times \gamma_{c|u})$ will be assigned to class c. As before, on average, $(\gamma_{c|u})$% of these individuals predicted to be in class c will be correctly classified, so the predicted misclassification rate with regard to class c will be $1 - \gamma_{c|u}{}^2$. Averaging over all c classes, the misclassification rate for individuals with this response pattern will be $\left(1 - \sum_{c=1}^{C} \gamma_{c|u}^2\right)$%, which will always be greater than the misclassification rate of $(1 - \gamma_{m|u})$% for modal assignment (Goodman, 2007).

Numerous factors influence the degree of misclassification in latent class models, the most significant of which is the degree of separation between latent classes. Good class separation results when response patterns have large posterior probabilities for the true latent class and near-zero posterior probabilities for all other classes, indicating little uncertainty about class membership. Uncertainty in class membership for a given response pattern manifests itself as similar posterior probabilities across 2 or more classes. The strength of association between the indicators and latent class membership is one factor that determines class separation: the more strongly indicators are associated with class membership (i.e., the more distinct conditional item probabilities are across classes), the greater the class separation (i.e., the more distinct posterior probabilities are across classes). Additionally, the number of indicators, number of classes and sample size can also influence class separation; generally, as the number of indicators increases, sample size increases, or the number of classes decreases, the less uncertainty in predicting class membership. Finally, the strength of association between the auxiliary variable and the latent class is also of concern when fitting latent regression models. The greater the strength of the association between an auxiliary variable and latent class, the more misclassification will bias the parameter estimates. For example, if there was truly no association between the auxiliary variable and latent class, misclassification would not bias estimates since individuals across all classes have the same value on the auxiliary variable.

All 3-step methods will underestimate parameter standard errors (SEs) since predicted class membership or estimated posterior probabilities are treated as fully observed, error-free variables (Bolck et al., 2004). Failure to account for the true uncertainty in this variable leads to underestimates of the SEs. The standard errors for multiple pseudoclass assignment will be larger than for modal assignment or posterior probability regression, given the increased variability across multiple draws, yet will still be artificially low, since the uncertainty of class membership is not reflected in the pseudoclass variable (Bolck et al., 2004).

Previous simulation studies have shown the misclassification arising from a 3-step approach using either modal or pseudoclass assignment yields significantly attenuated estimates of the parameters describing the association between the latent variable and the auxiliary variable (details discussed below). Additionally, given the greater misclassification under pseudoclass assignment, it has been shown that single and multiple pseudoclass assignment estimates are more attenuated relative to modal assignment estimates (Bolck et al., 2004; Clark & Muthén, 2009; Vermunt, 2010).

3.8 Corrected 3-step methods

Bolck, Croon, and Hagenaars (BCH) correction: Expanding on work by Croon (2002), Bolck et al. (2004) proved that the estimated odds ratios of class membership and an auxiliary variable in latent class regression using a classical 3-step approach will always be biased towards the null compared to estimates from a 1-step method. Furthermore, they quantified this bias, based on the general forms of the likelihood for the 1-step approach $\Pr(X, C, \boldsymbol{U}) = \Pr(X)\Pr(C|X)\Pr(\boldsymbol{U}|C)$ and for the 3-step approach $\Pr(X, C, \boldsymbol{U}, T) = \Pr(X)\Pr(C|X)\Pr(\boldsymbol{U}|C)\Pr(T|\boldsymbol{U})$, where *C* denotes true latent class, *X* denotes the covariate of interest, $\boldsymbol{U}$ denotes the array of latent class indicators, and *T* denotes predicted latent class. They showed the following relationship between the true quantity of interest, the distribution of the true latent class *C* given the observed covariate *X*, $\Pr(C|X)$, and it's 3-step proxy, $\Pr(T|X)$:

$$\Pr(T|X) = \sum_C \sum_u \Pr(T|\boldsymbol{U}) \Pr(\boldsymbol{U}|C) \Pr(C|X) = \sum_C \Pr(T|C) \Pr(C|X).$$

From this equation we see that $\Pr(T|X)$ is an average across all classes of $\Pr(C|X)$, weighted by the probability of misclassification, $\Pr(T|C)$. Theoretically, if there was no misclassification in the predicted latent classes, $\Pr(T|C) = 1$ and $\Pr(T|X)$ would be an unbiased estimate of $\Pr(C|X)$, yet due to the uncertainty inherent in the nature of unobserved latent variables, in practice $\Pr(T|C) \neq 1$.

The main idea of the Bolck et al. (2004) correction is to obtain a more accurate estimate for $\Pr(C|X)$ by correcting for this misclassification probability weighting in the estimation of $\Pr(T|X)$. Specifically, let *A* denote the matrix of $\Pr(C|X)$, our true quantity of interest; let *E* denote the matrix of $\Pr(T|X)$; and let *D* denote the matrix of misclassification probabilities $\Pr(T|C)$. The elements of *D* are of the form $d_{tc} = \Pr(T = t|C = c) = \sum_U \Pr(U|C) \Pr(T|U)$. The LCA results from Step 1 yield the posterior probabilities $\Pr(U|C)$ and the assignment rule (modal or pseudoclass) is represented by $\Pr(T|\boldsymbol{U})$, thus *D* can be calculated. Similarly, the elements of *E* are of the form $e_{tx} = \Pr(T = t|X = x) = \sum_U \Pr(T|\boldsymbol{U}) \Pr(\boldsymbol{U}|X)$. When *X* is categorical, $\Pr(\boldsymbol{U}|X)$ can be calculated as the proportion of individuals with X=x who have response pattern $\boldsymbol{U}$=$\boldsymbol{u}$ on the latent class indicators.

Bolck et al. (2004) show that *A* and *E* are related by the equation $A = ED^{-1}$. Thus, 3-step estimation can be used to generate $\Pr(T|X)$, which can then be corrected by multiplying these estimates by the inverse of the matrix *D*. When *D* is nonsingular, they show that this correction procedure generates consistent estimates of $\Pr(C|X)$.

Vermunt correction: Vermunt (2010) proposed an alternative correction and showed it to be more efficient than the BCH correction. The first two steps are the classical 3-step approach:

first, fit an LCA measurement model and, second, obtain predicted class membership (based on either modal or random assignment). Step 3 of Vermunt's method involves fitting another latent class model, including the auxiliary variable, using the predicted latent class variable as a single class indicator. To account for misclassification, the threshold probabilities within each of the *c* classes are constrained to be a function of the misclassification probability $\Pr(T|C)$. Specifically, the probability that an individual's true latent class $C = c_2$ given that his or her predicted latent class $T = c_1$ is calculated as follows:

$$p_{c1,c2} = \Pr(C = c_2|T = c_1) = \frac{1}{N_{c1}} \sum_{S_i = c1} \Pr\left(C = c_2 \,|\boldsymbol{U}\right),$$

where N_{c1} denotes the number of individuals predicted to be in class c_1. Then the misclassification probability is calculated as follows:

$$q_{c1,c2} = \Pr(T = c_1|C = c_2) = \frac{p_{c1,c2} N_{c1}}{\sum_c p_{c,c2} N_c}$$

In the LCA model in Step 3, the threshold for the indicator in class *c* is constrained to be $log\left(\frac{q_{c,c2}}{q_{c*,c2}}\right)$ where c_* denotes the reference latent class. See more details in Vermunt (2010) and Asparouhov & Muthén, (2013).

The model in Step 3 of this Vermunt correction is of the form $\Pr(T|X) = \sum_C \Pr(C|X) \Pr(T|C)$; maximum likelihood estimates can be obtained through any standard latent class software that allows fitting of a constrained model and polytomous indicators. Vermunt (2010) compare simulation results for a 1-step method and pseudo and modal assignment, both uncorrected and with his proposed correction, and demonstrates that his approach yields minimal bias and is nearly as efficient as 1-step methods. Mplus v7.11 now includes the option to use Vermunt's correction when estimating latent structural models with auxiliary variables (Asparouhov & Muthén, 2013).

Additionally, Vermunt highlighted the limitations of the BCH correction, including that it required categorical covariates, involved cumbersome matrix multiplication, and does not address the issue of underestimation of the parameter standard errors. He gives an alternate formulation of the BCH correction that does not require categorical covariates and shows that estimation of parameter standard errors using robust variance estimation corrects the attenuation in the SE estimates due to treating predicted latent class as an error-free variable. However, this improvement on the BCH correction is less efficient than Vermunt's new method.

Bakk et al. (2013) performed simulation studies examining the performance of Vermunt's correction method when regressing a distal outcome on latent classes and found that the Vermunt correction implemented using random pseudoclass assignment outperforms implementing it using modal assignment, particularly in the case of low class separation or weak association between the latent variable and the auxiliary variable.

3.9 Simulation Study: Methods

We conducted a simulation study to compare a 1-step method, three classical 3-step methods, and Vermunt's corrected 3-step method for class differences with respect to a distal outcome. The primary estimates of interest were the differences (i.e., contrasts) in distal outcome means across classes. Data was generated in Mplus v7.11 with a 3-class structure, 15 binary class indicators and a single normally distributed outcome Y (mean of Y varied across classes; variance of Y was constant across classes). Simulations investigated the effect of varying several aspects of the data: (1) class differentiation (i.e. entropy), (2) the magnitude of the differences in outcome across classes (referred to as "contrast size"), and (3) the variance (i.e. random noise) in Y. Latent class indicators were generated such that within a given class, all indicators had the same probability (conceptually, "low," "medium," or "high"). To vary class differentiation we considered the following sets of class-specific indicator probabilities (denoted p_1, p_2, p_3): (5%, 50%, 95%), (10%, 50%, 90%), (20%, 50%, 80%), (30%, 50%, 70%), which resulted in latent class models with mean entropies of 0.96, 0.90, 0.71, and 0.51 respectively. We varied the class specific means of Y, denoted $(\overline{Y}_1, \overline{Y}_2, \overline{Y}_3)$, in order to vary the magnitudes of the true contrasts

($\Delta_{21} = \bar{Y}_2 - \bar{Y}_1, \Delta_{31} = \bar{Y}_3 - \bar{Y}_1, \Delta_{32} = \bar{Y}_3 - \bar{Y}_2$). We considered the following 4 conditions: "Small" contrast magnitudes: $(\Delta_{21}, \Delta_{31}, \Delta_{32})$=(0.25, 0.50, 0.25); "Medium" contrast magnitudes: $(\Delta_{21}, \Delta_{31}, \Delta_{32})$ = (0.50, 1, 0.50); "Large" contrast magnitudes: $(\Delta_{21}, \Delta_{31}, \Delta_{32})$ = (1, 2, 1); and "Very Large" contrast magnitudes: $(\Delta_{21}, \Delta_{31}, \Delta_{32})$ = (1.5, 3, 1.5). Another factor affecting the strength of the association between latent class and outcome is the degree of variability, or random noise, in Y. We considered three conditions, SD(Y) = 0.5, SD(Y) = 1, and SD(Y) = 2. Each simulated dataset contained 5,000 observations and 1,000 simulations were performed at each setting.

The 1-step method was implemented in R using the LCCA (Latent Class Causal Analysis) package by Kang and Schafer (2010) using the *lcca* function. We implemented classical 3-step methods – modal assignment, 20 pseudoclass draws, and posterior probability regression – in R. Finally, we implemented Vermunt's correction method, based on modal assignment, in Mplus v7.11 using the new Auxiliary = DU3STEP option for mixture models. We report bias, standard error (SE), mean squared error (MSE), and the 95% confidence interval coverage rate, namely the percentage of 95% confidence intervals that contained the true difference in means. (Note that 95% coverage estimates were not available in Mplus). To facilitate comparison across the 3 contrasts of interest, we standardized bias estimates for each contrast by the magnitude of the corresponding true contrast.

3.10 Simulation Study: Results

Figure 3.5 and Table 3.1 presents bias, SE, MSE, and 95% confidence coverage rates for the Class 2 – Class 1 contrast under all 5 methods when both entropy and effect size vary. We also estimated the Class 3 – Class 1 and the Class 3 – Class 2 contrasts; the trends were the same except where discussed below. We considered four entropy levels and four sets of contrast magnitudes, for a total of sixteen different conditions. The Bias panel in Figure 5 shows that the 1-step method, as expected, has near-zero bias under all conditions; that is, neither entropy nor the contrast magnitudes markedly affects bias. As shown in Table 3.1, the mean bias is positive for 9 out of the 16 conditions and was negative for the remaining 7, which is consistent with previous simulation studies that have demonstrated that 1-step methods have a slight positive bias (Clark & Muthén, 2009; Vermunt, 2010). Modal and pseudoclass assignment both yield significant biases and show the same trends – the magnitude of bias increases as entropy decreases (misclassification increases), yet the magnitudes of true class differences do not affect bias. As expected, the average bias is negative for all estimated effects under all 16 conditions for both modal and pseudoclass assignment, indicating that these methods underestimate the true association of latent class and the distal outcome Y. Bias for pseudoclass assignment is consistently larger than that for modal assignment, due to the higher rates of misclassification under pseudoclass (as discussed previously). The bias for posterior probability regression and the Vermunt correction are similar in magnitude to bias for the 1-step method, and are often several orders of magnitude smaller than the bias for modal or pseudoclass assignment. Similar to the 1-step method, the bias for posterior probability regression and the Vermunt correction are both positive and negative, although for the Class 2 – Class 1 contrast they are predominately negative.

Figure 3.5 shows that the standard errors estimated under the 1-step method are larger than under any other method, highlighting that 3-step methods, particularly modal and pseudoclass assignment, attenuate SE estimates. For all methods, the SEs increase as entropy decreases, although this trend is less pronounced for modal and pseudoclass assignment. For modal and pseudoclass assignment, SEs increase as contrast size increases; there is not a clear association between contrast size and SEs for the other three methods. The 1-step method, probability regression, and the Vermunt correction all have small mean squared errors, particularly when entropy is high. Since these three methods had very little bias, MSE is largely driven by trends in variance rather than bias. Mean squared error for modal and pseudoclass assignment increases both as entropy decreases and contrast size increases. Finally, Figure 3.5 highlights the limitations of modal and pseudoclass assignment with respect to 95% CI coverage–

while these methods can achieve nominal coverage with high entropy and small contrast sizes, they show very poor coverage as entropy decreases and the contrast magnitude increases. Posterior probability regression shows the same trend of decreasing coverage with decreasing entropy or increasing class difference magnitude, yet achieves nominal coverage in 10 of 16 conditions. Coverage rates for the 1-step method are excellent and are not affected by either entropy or the magnitude of class contrasts.

Figure 3.6 and Table 3.2 presents bias, SE, MSE, and 95% confidence coverage rates for the Class 2 – Class 1 contrast under all 5 methods when both the variance (i.e., random noise) of the distal outcome Y and contrast size varies. The Bias panel of Figure 3.6 shows that the 1-step, Vermunt correction, and probability regression methods yield bias several orders of magnitude smaller than the bias from modal and pseudoclass assignment. As before, bias for all contrasts are negative for modal and pseudoclass assignment; for the other three methods, bias for the Class 2 – Class 1 contrast are negative but bias for the other two contrasts are both positive and negative (full results not shown). Variance of the distal outcome Y does not affect bias for any method, nor does contrast size. Additionally, Figure 3.6 highlights that standard error magnitude is associated with the variance in Y for all methods. The standard errors for modal and pseudoclass assignment are consistently smaller than the standard errors for the 1-step methods, as in the previous results. Mean squared error increases with both increasing variance in Y and contrast size for modal and pseudoclass assignment, but MSE is only affected by the magnitude of variance in Y for the other three methods. Finally, the 1-step method and probability regression both show excellent 95% confidence coverage rates that are not affected by contrast size or variance in Y. On the other hand, coverage for modal and pseudoclass assignment notably decreases with increasing contrast size and decreasing variance in Y.

3.11 Discussion

This paper provides an overview of 1-step and 3-step methods for latent variable regression for applied researchers and describes current methods for correcting classical 3-step methods. We present results from a simulation study comparing the performance of 1-step, classical 3-step and corrected 3-step methods for estimating the association between a latent variable and a continuous distal outcome. Our results highlight that 1-step methods, such as the 1-step method implemented in the LCCA package in R, perform very well under a wide range of conditions. Specifically, neither the degree of class separation (i.e., entropy), the magnitude of outcome differences across class (i.e., contrast size), nor the random variance in the distal outcome significantly affect bias, mean squared error, or 95% confidence interval coverage of the estimated associations between latent class and outcome. The Vermunt correction, as implemented in Mplus v7.11 with modal assignment, also performs very well with regard to bias and mean squared error (95% coverage rates not available), although standard error estimates are somewhat attenuated relative to the 1-step method. Of the classical three-step methods, posterior probability regression performs the best, showing little bias, unlike either modal or pseudoclass assignment. Probability regression performs similar to the 1-step method in terms of standard error and 95% coverage rates when classes are reasonably well separated (entropy > 0.70), yet performance declines with decreasing entropy. Finally, modal and pseudoclass assignment yield less than ideal performance, with significant bias and mean squared error, as well as markedly underestimated standard errors and poor 95% coverage rates. Poor class separation affects the performance of pseudoclass assignment more than modal assignment, resulting in greater bias and mean squared error and lower 95% coverage rates.

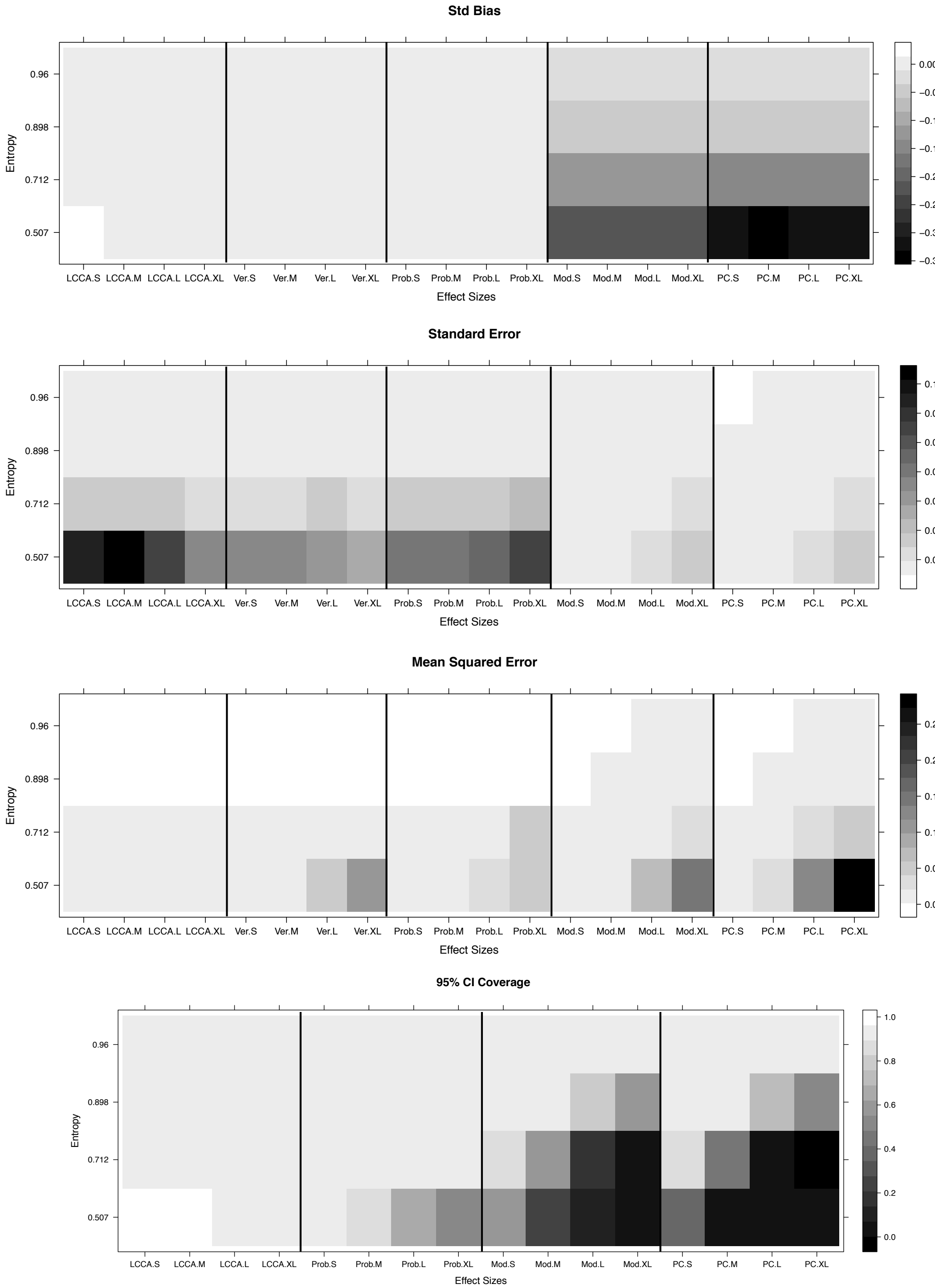
Although many estimation issues are similar for latent class regression with a predictor and with a distal outcome and previous work indicates that statistical methods perform similarly for these two types of latent class regression (Asparouhov & Muthén, 2013, Bolck, et al., 2004;

Vermunt, 2010), the statistical considerations for each method are not identical. Thus, our simulation results may not be fully generalizable to the latent class regression with a predictor. One notable difference between regression with a predictor and with a distal outcome is that while 1-step methods are typically conceptually sound when modeling a predictor, they may not be as conceptually appropriate when modeling a distal outcome. Specifically, when one is interested in estimating causal effects of the latent class on the distal outcome, it is not desirable to let the outcome influence the estimation of the latent classes, which is possible under joint modeling. Thus, corrected 3-step methods may be especially important for the case of distal outcomes in order to separate estimation of the measurement and structural models. One can compare the latent classes estimated with the joint model to the latent classes estimated with an unconditional LCA model (i.e., one that includes no auxiliary variables) in order to assess the influence of the auxiliary variable on latent class formation; if the definition of the latent classes changes markedly, a corrected 3-step approach may be advisable.

One limitation to our simulation study was that the measurement and structural models were correctly specified in all analyses. In practice, either may be misspecified; for example, the measurement model may misspecify the true number of latent classes or the structural model may misspecify the nature of the relationship between the latent variable and the auxiliary variable. Our simulations did not investigate how robust these various methods are to model misspecification; it is likely that model misspecification could significantly diminish their performance, perhaps differentially. As work by Sanchez et al. (2009) suggests, 1-step methods may be somewhat more sensitive to model misspecification than 3-step methods, since joint modeling assumes that both the measurement and the structural model are correctly specified. Future work should investigate the robustness of these methods to model misspecification and other conditions, in order to elucidate any conditions under which corrected 3-step methods outperform 1-step methods.

In general, latent variable regression represents a powerful statistical method that is gaining popularity in the social, behavioral, and medical fields. This paper is intended to highlight the pitfalls of classical 3-step methods for latent class regression and to promote the use of 1-step methods or corrected 3-step methods. One-step estimation for auxiliary predictors is available in many standard statistical packages, including Mplus, Latent Gold, various LCA packages in R, SAS PROC LCA. One-step estimation for distal outcomes is increasingly available as well. As discussed, 1-step estimation may not be computationally feasible or conceptually appropriate for all data; it is essential that applied researchers also have corrected 3-step methods in their statistical repertoire. In recent years, several corrected 3-step methods have been developed and these methods are increasingly available in standard statistical packages. Thus, when 3-step methods are more appropriate, should use corrected 3-step methods rather than classical 3-step methods.

Figure 3.5. Standardized bias, standard error, mean squared error and 95% confidence interval coverage as a function of both contrast size and entropy for the Class 2 – Class 1 contrast.



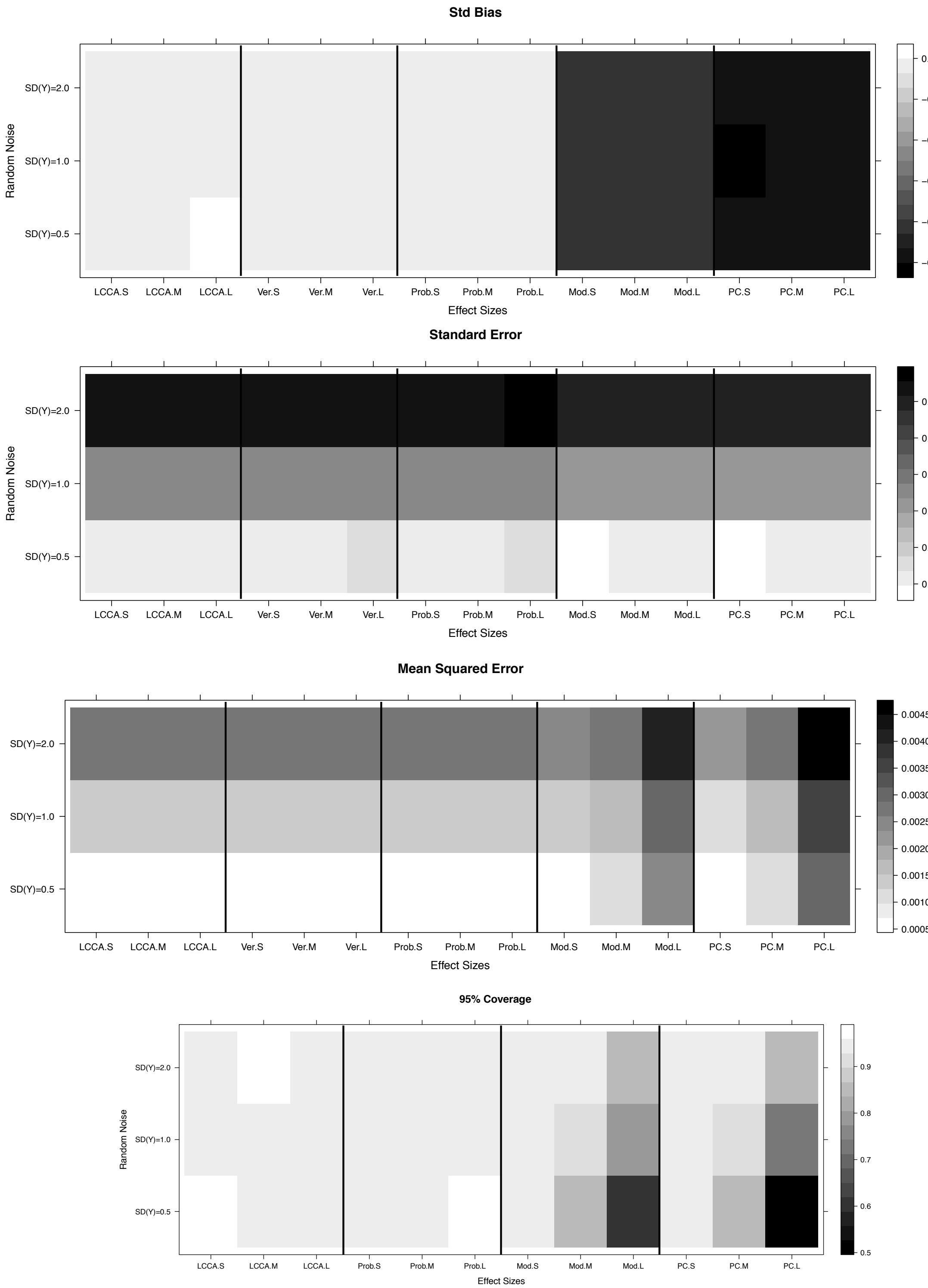
Abbreviations: LCCA = 1-step method; Ver = Vermunt correction; Prob = posterior probability regression, Mod = Modal assignment; PC = pseudoclass assignment. S, M, L and XL suffixes denote contrast sizes of 0.25, 0.50, 1.0, and 1.5, respectively. In all figures dark shading indicates worse performance.

Table 3.1. Standardized bias, standard error, mean squared error and 95% confidence interval coverage as a function of both contrast size and entropy for the Class 2 – Class 1 contrast.

	Contrast Size: C2 v C1 = 0.25				*Contrast Size: C2 v C1 = 0.50*			
	E=0.95	*E=0.90*	*E=0.70*	*E=0.50*	*E=0.95*	*E=0.90*	*E=0.70*	*E=0.50*
				Bias / Δ				
1-step	-0.001	-0.001	0.004	0.015	0.000	0.000	0.002	0.010
3-step Modal	-0.018	-0.042	-0.117	-0.231	-0.018	-0.042	-0.119	-0.232
3-step Pseudoclass	-0.019	-0.050	-0.149	-0.331	-0.019	-0.050	-0.151	-0.332
Prob Regression	-0.001	-0.001	0.002	-0.001	-0.001	-0.001	0.000	-0.001
Vermunt correction	0.000	-0.003	-0.001	-0.001	0.000	-0.002	-0.003	-0.002
				Standard Error				
1-step	0.036	0.037	0.045	0.096	0.035	0.037	0.045	0.102
3-step Modal	0.035	0.035	0.035	0.036	0.035	0.035	0.036	0.037
3-step Pseudoclass	0.035	0.035	0.035	0.036	0.035	0.035	0.036	0.038
Prob Regression	0.035	0.037	0.045	0.069	0.036	0.037	0.045	0.071
Vermunt correction	0.035	0.037	0.043	0.066	0.035	0.037	0.044	0.067
				Mean Squared Error				
1-step	0.0012	0.0014	0.0019	0.0070	0.0012	0.0013	0.0019	0.0071
3-step Modal	0.0012	0.0013	0.0021	0.0055	0.0012	0.0016	0.0050	0.0188
3-step Pseudoclass	0.0011	0.0012	0.0021	0.0081	0.0012	0.0016	0.0065	0.0313
Prob Regression	0.0012	0.0014	0.0019	0.0058	0.0012	0.0014	0.0019	0.0098
Vermunt correction	0.0012	0.0013	0.0018	0.0078	0.0012	0.0014	0.0020	0.0179
				95% CI Coverage				
1-step	95.4%	95.9%	95.5%	96.4%	95.3%	95.7%	95.2%	96.4%
3-step Modal	95.2%	94.6%	87.6%	61.8%	95.0%	91.4%	60.4%	23.7%
3-step Pseudoclass	95.7%	95.0%	88.0%	34.8%	94.8%	91.1%	42.1%	5.4%
Prob Regression	95.4%	95.7%	95.5%	93.0%	95.6%	95.8%	95.5%	83.0%
	Contrast Size: C2 v C1 = 1.0				*Contrast Size: C2 v C1 = 1.5*			
	E=0.95	*E=0.90*	*E=0.70*	*E=0.50*	*E=0.95*	*E=0.90*	*E=0.70*	*E=0.50*
				Bias / Δ				
1-step	0.000	0.000	0.001	0.007	0.000	0.000	0.000	0.003
3-step Modal	-0.017	-0.042	-0.120	-0.230	-0.017	-0.042	-0.120	-0.229
3-step Pseudoclass	-0.019	-0.049	-0.152	-0.331	-0.018	-0.049	-0.152	-0.330
Prob Regression	0.000	-0.001	-0.001	-0.002	0.000	-0.001	-0.002	-0.001
Vermunt correction	0.000	-0.002	-0.003	-0.003	0.000	-0.002	-0.003	-0.005
				Standard Error				
1-step	0.035	0.037	0.044	0.086	0.035	0.037	0.042	0.065
3-step Modal	0.035	0.036	0.037	0.040	0.035	0.037	0.040	0.045
3-step Pseudoclass	0.035	0.036	0.038	0.042	0.036	0.037	0.042	0.049
Prob Regression	0.036	0.038	0.047	0.076	0.036	0.038	0.049	0.084
Vermunt correction	0.035	0.037	0.044	0.062	0.036	0.038	0.044	0.057
				Mean Squared Error				
1-step	0.0012	0.0013	0.0018	0.0059	0.0012	0.0013	0.0017	0.0040
3-step Modal	0.0015	0.0030	0.0166	0.0692	0.0019	0.0054	0.0362	0.1512
3-step Pseudoclass	0.0015	0.0035	0.0243	0.1227	0.0019	0.0066	0.0538	0.2730
Prob Regression	0.0012	0.0014	0.0022	0.0277	0.0012	0.0014	0.0027	0.0550
Vermunt correction	0.0012	0.0014	0.0025	0.0550	0.0012	0.0014	0.0030	0.1083
				95% CI Coverage				
1-step	95.4%	95.8%	95.5%	96.0%	95.3%	94.7%	95.4%	96.0%
3-step Modal	93.3%	77.6%	16.8%	9.2%	90.5%	58.1%	6.0%	5.8%
Prob Regression	93.2%	74.1%	1.7%	1.4%	90.2%	48.3%	0.0%	0.7%
Prob Regression	95.7%	95.8%	94.3%	65.2%	95.6%	95.8%	93.7%	52.7%

Abbreviations: E = entropy, Bias / Δ = bias standardized by the true contrast size

Figure 3.6. Bias, standard error, mean squared error and 95% confidence interval coverage as a function of both contrast size and variance in the distal outcome for the Class 2 – Class 1 contrast.



Abbreviations: LCCA = 1-step method; Ver = Vermunt correction; Prob = posterior probability regression, Mod = Modal assignment; PC = pseudoclass assignment. S, M, L and XL suffixes denote contrast sizes of 0.25, 0.50, 1.0, and 1.5, respectively. In all figures dark shading indicates worse performance.

Table 3.2. Standardized bias, standard error, mean squared error and 95% confidence interval coverage as a function of both contrast size and variance in the distal outcome for the Class 2 – Class 1 contrast.

	Contrast: C2 v C1=0.25			*Contrast: C2 v C1=0.50*			*Contrast: C2 v C1 =1.0*		
	SD(Y) = 0.5	*SD(Y) = 1.0*	*SD(Y) = 2.0*	*SD(Y) = 0.5*	*SD(Y) = 1.0*	*SD(Y) = 2.0*	*SD(Y) = 0.5*	*SD(Y) = 1.0*	*SD(Y) = 2.0*
					Bias / Δ				
1-step	0.000	-0.001	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
3-step Modal	-0.042	-0.042	-0.043	-0.042	-0.042	-0.042	-0.042	-0.042	-0.042
3-step Pseudoclass	-0.049	-0.050	-0.050	-0.050	-0.050	-0.049	-0.050	-0.049	-0.049
Prob Regression	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
Vermunt correction	-0.003	-0.003	-0.004	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002
					Standard Error				
1-step	0.026	0.037	0.052	0.026	0.037	0.052	0.026	0.037	0.052
3-step Modal	0.025	0.035	0.049	0.025	0.035	0.049	0.026	0.036	0.050
3-step Pseudoclass	0.025	0.035	0.049	0.025	0.035	0.050	0.026	0.036	0.050
Prob Regression	0.026	0.037	0.052	0.026	0.037	0.052	0.027	0.038	0.053
Vermunt correction	0.026	0.037	0.052	0.026	0.037	0.052	0.027	0.037	0.052
					Mean Squared Error				
1-step	0.0007	0.0014	0.0027	0.0007	0.0013	0.0027	0.0006	0.0013	0.0027
3-step Modal	0.0007	0.0013	0.0025	0.0011	0.0016	0.0028	0.0025	0.0030	0.0042
3-step Pseudoclass	0.0007	0.0012	0.0022	0.0011	0.0016	0.0027	0.0030	0.0035	0.0045
Prob Regression	0.0007	0.0014	0.0027	0.0007	0.0014	0.0027	0.0007	0.0014	0.0027
Vermunt correction	0.0007	0.0013	0.0027	0.0007	0.0014	0.0027	0.0007	0.0014	0.0027
					95% CI Coverage				
1-step	96.0%	95.9%	95.8%	95.8%	95.7%	96.0%	94.8%	95.8%	95.8%
3-step Modal	93.5%	94.6%	95.0%	86.4%	91.4%	93.6%	60.5%	77.6%	86.4%
3-step Pseudoclass	94.8%	95.0%	95.9%	84.2%	91.1%	94.7%	52.6%	74.1%	85.2%
Prob Regression	95.8%	95.7%	95.7%	95.9%	95.8%	95.8%	96.0%	95.8%	95.9%

Abbreviations: Bias / Δ = bias standardized by the true contrast size

CHAPTER 4. ADDRESSING CONFOUNDING WHEN ESTIMATING THE EFFECTS OF LATENT CLASSES ON A DISTAL OUTCOME

4.1 Abstract

Background: Although confounding is widely recognized in settings where all variables are fully observed, recognition of and statistical methods to address confounding in the context of latent variable regression are slowly emerging. Although standard methods of addressing confounding can be extended to latent variable regression, this is not as straightforward when the treatment of interest is a latent variable. This study describes a recent 1-step method for addressing confounding when regressing a distal outcome on latent classes, as well as an approach for incorporating propensity score weighting with two different 3-step methods; we compare the statistical performance of these three methods on simulated data. Additionally, these methods are applied to data on adolescent substance use treatment in order to examine the effect of classes of drug treatment services on substance use outcomes.

Methods: In Part I, we perform a simulation study to compare a 1-step method, modal assignment with propensity score weighting, and multiple pseudoclass assignment with propensity score weighting. Data were generated with 3-class structure, 15 binary class indicators, 8 normally distributed covariates (i.e., confounders), and a continuous outcome variable; we considered conditions with varying degrees of class separation (i.e., entropy) and magnitude of confounding. All analyses were performed in R; the 1-step method used the LCCA package. We also implemented the 1-step method without covariate adjustment and modal assignment without propensity score weighting to assess the impact of ignoring potential confounding. The estimates of interest were the pairwise differences in outcome means across classes; statistical performance was assessed in terms of bias, standard error, mean squared error, and 95% confidence interval coverage. In Part II, we used data on 5,527 adolescents receiving outpatient drug treatment services from treatment providers throughout the US who were funded by the Substance Abuse and Mental Health Service Administration's Center for Substance Abuse Treatment. Latent classes of drug treatment services were estimated based on 12 items from the Global Appraisal of Individual Needs' Treatment Received Scale. The distal outcome of interest was the Substance Problems Scale.

Results: In our simulation studies, we found that the 1-step method, modal assignment with propensity score weighting, and pseudoclass assignment with propensity score weighting all significantly reduced the bias arising from confounding relative to the unadjusted 1-step and modal assignment approaches. However, the 1-step method performed better with regard to bias and 95% confidence interval coverage, particularly when class separation was poor (i.e., low entropy). Our applied example also highlighted the importance of addressing confounding – both unadjusted methods indicated significant differences in substance use problems across treatment classes, yet these class differences were not significant under any of the three adjusted methods.

Conclusion: Potential confounding should be carefully considered when conducting latent variable regression with a distal outcome; failure to do so may results in significantly biased effect estimates or spurious statistical inferences. Currently statistical methods to address confounding in this context are limited; future work should continue to develop and refine statistical methods to address confounding when the treatment of interest is a latent variable.

4.2 Introduction

Latent variable modeling is an increasingly popular statistical method in public health and social science since many constructs of interest in these fields are not directly observable. For example, mental health conditions, such as depression, are not directly observable but rather measured through a diagnostic checklist. Standard analysis approaches would treat depression status, as measured by these diagnostic items, the same as any fully observed variable, such as gender or clinic site. On the other hand, latent variable methods appropriately account for the measurement error inherent in using a set of observed items to represent an underlying latent construct, resulting in more appropriate statistical inferences.

One common type of latent variable modeling is latent variable regression, which models the association between a latent variable and auxiliary variables of interest (either predictors or distal outcomes). Latent variable regression is frequently conducted on observational data, yet potential confounding in latent variable regression is rarely addressed, despite the fact that applied researchers are increasingly adopting advanced techniques to address confounding, such as propensity score methods. Recent work by Lanza et al. (2013a) and Butera et al. (2013) has proposed propensity score based methods to address confounding in latent class regression when latent class is regressed on an observed predictor (or "treatment") of interest. However, it is less straightforward to extend propensity score methods when regressing a distal outcome on a latent predictor, since the treatment of interest is now latent, rather than observed. Analyses that do not adjust for baseline differences across the latent classes will conflate the true effect of latent class membership on the distal outcome with preexisting group differences. Thus, latent variable regression requires as much critical attention to confounding as analyses of observational data with fully observed variables.

The motivating example for this paper involves estimating treatment effects of substance use treatment services on substance use outcomes for adolescents. Using national evaluation data from outpatient treatment providers funded through SAMHSA's Center for Substance Abuse Treatment (CSAT), we estimated latent classes of treatment services received by youth in the typical course of outpatient substance treatment (see Chapter 2). Latent classes were formed based on twelve items from the Global Assessment of Individual Needs (GAIN) survey's Treatment Received Scale (Dennis, 1999), which asks about services such as biological drug testing, family involvement, and case management services. We identified four latent classes described as Class 1: Low Service Utilization; Class 2: Individual-Focused Services; Class 3: Individual- and Family-Focused Services; and Class 4: Multiple Services. Having identified these latent treatment classes, our goal is to estimate the causal effects of these classes on severity of substance use problems.

Given the observational nature of our data, it will be important to control for baseline differences across groups when estimating the effect of treatment class on substance use outcomes. We found that class membership was associated with demographic characteristics, justice system involvement and baseline substance use. Since it is plausible that these factors, particularly baseline substance use, may be associated with subsequent substance use outcomes, failing to account for baseline differences could lead to biased effect estimates, just as is the case with non-experimental studies more generally. An unadjusted analysis shows that youth in Class 4, who are receiving the largest number of different services, have the highest mean score on the GAIN's Substance Problems Scale (Past Month) at 3 months, suggesting that these youth have the most substance problems at 3 months. However, this group also exhibits the highest mean score on the Substance Problems Scale at baseline. Thus, we want to appropriately control for the observed baseline differences when comparing the effects of our latent classes of treatment services in order to obtain unbiased estimates of effects.

In this paper, we first provide a discussion of the challenges associated with addressing confounding when estimating the effect of a latent variable on a distal outcome and review current methods. We then conduct a simulation study that compares three proposed methods for addressing confounding, as well as two methods that do not adjust for potential confounding in order to emphasize the potential for bias when confounding is not addressed. Finally, we apply these methods to our adolescent substance treatment dataset in order to address the substantive question at hand. We highlight that the statistical inference can change dramatically when confounding is not addressed.

4.3 Latent Class Analysis

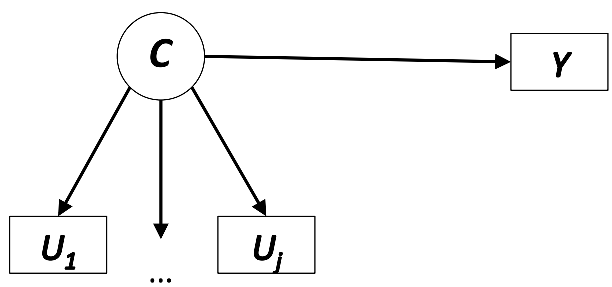
Latent class analysis (LCA) is a widely used latent variable model that models the latent variable as categorical, that is, assumes an underlying structure of discrete classes (denotes as C). Each individual belongs to exactly one latent class. An individual's latent class membership is determined from his or her observed response patterns across multiple indicator items $U_1, U_2, \ldots, U_j$ (Collins & Lanza, 2010; Goodman, 1974; Haberman, 1979; Hagenaars & McCutcheon, 2002; Lazarsfeld & Henry, 1968).

The two parameters of primary interest in a latent class analysis are the posterior class membership probabilities and the conditional item probabilities. The posterior class membership probabilities represent an individual's likelihood of belonging to class c, given his or her observed response pattern $\boldsymbol{u}$, and are denoted $\gamma_{c|u} = \Pr(C = c|\boldsymbol{U} = \boldsymbol{u})$. The conditional item probabilities represent the probability that an individual in latent class c will answer indicator U_j with response r_j, and are denoted $\rho_{j,rj|c} = \Pr\left(U_j = r_j \mid C = c\right)$, and are used to interpret the meaning of each latent class. Classically, the latent class analysis model is expressed as follows: $\Pr(\boldsymbol{U} = \boldsymbol{u}) = \sum_{c=\mathbf{1}}^{C} \Pr(C = c) \prod_{j=\mathbf{1}}^{J} \Pr\left(U_j = r_j \mid C = c\right)$. A fundamental assumption of LCA is *local independence*, which states that the indicator items $U_1, U_2, \ldots, U_j$ are mutually independent after conditioning on latent treatment class membership C.

4.4 Latent Class Analysis with Distal Outcomes

Latent variable models that regress latent class on predictive covariates have long been used in social and behavioral research and been available standard in latent variable statistical software; methods to estimate the opposite association, that is regress a distal outcome on latent class, have been developed and adopted more recently and are the focus of this paper (see Figure 4.1). Like standard LCA models, LCA with distal outcomes requires the assumption that the indicators are mutually independent, given class membership. Additionally, as discussed by Lanza et al. (2013b), it assumes conditional independence of the distal outcome *Y* and the indicators.

Figure 4.1. Schematic figure of latent class analysis with distal outcomes. $U_1 – U_j$ denote latent class indicators.



Latent class analysis with distal outcomes may be conducted using either a 1-step or 3-step method. The relative merits of each approach have been previously discussed (see Bolck et al., 2004; Feingold et al., 2013; Vermunt, 2010; see Chapter 3). In brief, 1-step methods fit a joint model that simultaneously estimates the latent class measurement model and the structural model describing the relationship between the latent variable and the auxiliary variable. In general, 1-step methods yield unbiased and efficient parameter estimates, yet may not converge in some cases due to complexity of the joint likelihood, are not easily implemented for all possible analyses, and are not available in all statistical packages. Thus, 3-step methods are also commonly used, the most common of which are modal assignment and pseudoclass assignment. Three-step methods first fit an unconditional (i.e., no covariates) latent class model and then predict latent class membership based on the estimated posterior probabilities. Then, the association between latent class and auxiliary variable is estimated through a regression model using predicted latent class. Under modal assignment, individuals are predicted to be in the latent class for which they have the highest posterior probability (Clogg, 1995). Under pseudoclass assignment, latent class membership is predicted by random draws from a multinomial distribution defined by an individual's posterior probabilities (Bandeen-Roche et al., 1997; Goodman, 2007; Wang et al., 2005); pseudoclass assignment is often performed multiple times (e.g., 20) and results across draws are combined using combining rules for multiple imputation (Rubin, 1987).

4.5 Propensity Score Methods as a Means to Address Confounding

Propensity score methods are a standard method for addressing selection bias in an observational study (Rosenbaum & Rubin, 1983; Rubin, 2001; Stuart, 2010). In the case of two treatment groups, the *propensity score* is defined as the probability that an individual received the treatment, conditional on the individual's observed covariates, and is denoted $p(x) = \Pr(T_i = t|\ \boldsymbol{X}_i = \boldsymbol{x})$ where $T = \{0, 1\}$ and $\boldsymbol{X}$ represents a vector of observed covariates. The propensity score can be extended to cases in which there are more than two treatment groups (as will often be the case when latent classes represent treatment groups); Imbens (2007) refers to this as the *generalized propensity score*, defined as $p(t, x) = \Pr(T_i = t|\ X_i = x)$ where $t \in \mathcal{T}$.

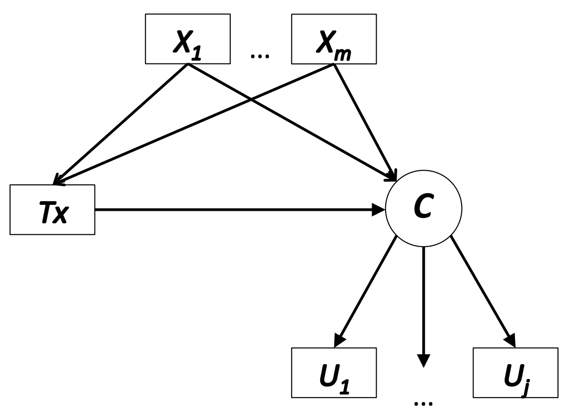
Propensity scores can be incorporated in the final analysis through propensity score matching, subclassification, or weighting; we primarily focus on propensity score weighting in this paper – details on other methods can be found in (Stuart, 2010). Propensity score weighting implements a weighted regression, in which each individual's weight is a function of his or her propensity score. A common weighting approach is Inverse Probability of Treatment Weighting (IPTW), which weights each individual by the inverse probability of receiving the treatment he or she truly did receive, essentially weighting both the treatment and control group to look like the overall study population. McCaffrey et al. (2013) describe an extension of IPTW for more than two treatment groups.

Propensity score methods are preferable to regression covariate adjustment for several reasons. First, propensity score methods do not necessarily rely on the parametric modeling assumptions required by regression adjustment (Ho et al., 2007). Additionally, propensity score methods avoid potential bias that arises from extrapolating beyond observed data in traditional regression models when the treatment groups have little overlap in terms of covariates (Stuart, 2010). Furthermore, propensity scores are an effective dimension reduction technique when there are a substantial number of baseline covariates to adjust for (Rosenbaum & Rubin, 1984). Finally, as advocated by Rubin, it is philosophically cleaner to separate the analytic step of controlling for confounding from the step of implementing the final structural model (Rubin, 2001). Separation prevents potential bias that may arise from adjusting for covariates solely because they favorably influence the treatment effect estimates.

4.6 Confounding in Latent Variable Regression when Treatment is Observed

One common type of latent variable regression model, often called latent class analysis with covariates, estimates the association between observed predictors and the latent variable. As recent work by Lanza et al. (2013a) and Butera et al. (2013) show, propensity score methods can be incorporated to balance groups on the observed predictor of interest in order to facilitate estimation of causal effects on the latent variable. The method they outline first fits a propensity score model estimating the association between the predictor of interest and potential confounders. After ensuring that the estimated propensity scores achieve good balance on the covariates, the propensity score method of interest can be implemented; Lanza et al. (2013) discuss propensity score matching and weighting. If matching, the estimated propensity scores are used to create a matched dataset in which the treated individuals are matched to control individuals who have similar propensity score values; if weighting, propensity score-based weights are calculated. Then the appropriate number of latent classes is determined by fitting a series of unconditional latent class models, incorporating the propensity score approach (e.g., by using the matched dataset or the propensity score-based weights). After determining the optimal number of classes, the final step fits a latent class regression model that includes the predictor as an auxiliary variable, again incorporating the propensity score approach. If adequate balance on baseline covariates is obtained through use of propensity scores, then the estimates of the association between the predictor and the latent variable obtained from the final model will appropriately control for potential confounding due to observed variables.

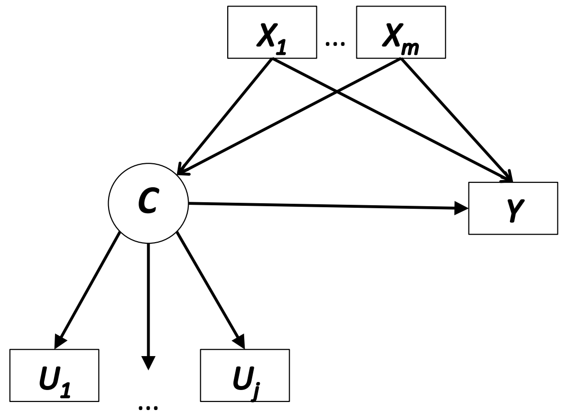
Figure 4.2. Schematic figure of latent variable regression with confounding when the treatment (Tx) is observed. $U_1 - U_j$ denote latent class indicators; $X_1 - X_m$ denote potential confounders.



4.7 Confounding in Latent Variable Regression when Treatment is Latent

Controlling for confounding in the context of latent class analysis with a distal outcome is less straightforward since the treatment (or exposure) variable of interest is defined as a latent variable. We now discuss various extensions of latent class analysis with distal outcomes that can account for multiple potential confounders (see Figure 4.3).

Figure 4.3. Schematic figure of latent variable regression with confounding when the treatment is a latent variable (C). $U_1 - U_j$ denote latent class indicators; $X_1 - X_m$ denote potential confounders.



A 1-step method can be used to jointly model the latent class indicators ***U***s, the potential confounders ***X***s, and the distal outcome *Y*; recently, Kang and Schafer (2010) proposed a 1-step method known as Latent Class Causal Analysis. In the LCCA model, the vector of parameters of interest are denoted $\theta = \{\rho, \alpha, \beta, \Sigma\}$, where ρ represents the LCA indicator-response probabilities, α are treatment model coefficients, β are outcome model coefficients, and Σ is the covariance matrix. LCCA specifies the following likelihood, which is maximized using the Expectation-Maximization (EM) algorithm:

$$l_i(\theta) = \ log \sum_{c=1}^{C} \left\{ \frac{\exp\,(x_i^T \alpha_c)}{\sum_{c'=1}^{C} \exp\,(x_i^T \alpha_{c\prime})} \right\} \left\{ \prod_{m \in obs_i} \prod_{r=1}^{r_m} \rho_{mr|c}^{I(u_{im}=r)} \right\} \left\{ (2\pi\sigma_c^2)^{-1/2} exp \left\{ \frac{-1}{2\sigma_c^2} \left(y_{i,obs} - x_i^T \beta_c \right)^2 \right\} \right\}$$

Estimates of the Average Treatment Effect (ATE) are then obtained from the maximum-likelihood parameter estimates via expected estimating equations (Wang et al., 2008). LCCA is implemented in the LCCA package for *R* (Kang & Schafer, 2010; Schafer & Kang, 2013).

This 1-step method for latent variable regression with confounders faces the same limitations as 1-step methods for latent variable regression. Particularly, the 1-step method may not converge under all conditions, given the added complexity of the joint model due to the inclusion of the confounders. Furthermore, these methods require specialized software in order to maximize the joint likelihood; although 1-step methods for latent variable regression are currently available in some statistical packages, including Mplus and SAS, 1-step methods that address confounding are not widely available. Given the implementation challenges of 1-step approaches, and the fact that they do not utilize propensity score methods, which may more flexible and yield better statistical performance in some settings, we investigate the incorporation of propensity score methods with two commonly used 3-step approaches to latent variable regression, namely modal assignment and pseudoclass assignment.

One potential benefit of 3-step approaches is that, because they essentially turn the latent treatment variable into an observed treatment variable, it is fairly straightforward to combine them with other common statistical procedures. Specifically, it is possible to combine propensity score methods with either modal or pseudoclass assignment for latent class analysis with distal outcomes by estimating a propensity score model based on the predicted latent class. Propensity scores can be incorporated in the final analysis through propensity score matching, subclassification, or weighting. Propensity score matching and subclassification are most easily applicable in the case of two treatment groups; propensity score weighting is fairly easily extended to settings where there are more than two treatment groups (see McCaffrey et al., 2013). Given that latent class analysis often produces more than two classes, as in our motivating example, propensity score weighting will likely be the easiest propensity score approach to implement when defining treatment groups based on latent classes. Thus, we primarily describe propensity score weighting, although matching or subclassification could alternatively be used in the case of two classes.

The general outline for incorporating propensity scores into a 3-step approach is as follows: (1) fit an unconditional latent class analysis model and obtain posterior probability estimates; (2) use either modal or pseudoclass assignment to create the predicted latent class variable; (3) estimate the propensity score model by regressing predicted latent class on potential confounders; (4) calculate propensity score weights and assess covariate balance after weighting; (5) fit a weighted regression model for the distal outcome on predicted latent class, using propensity score weights.

4.8 Part I: Simulation Study

Methods

First, we conducted a latent class simulation study to compare Schafer & Kang's 1-step method to the proposed 3-step approaches, modal assignment with propensity score weighting and pseudoclass assignment with propensity score weighting. In addition to these three methods, we also considered the 1-step method without covariate adjustment and modal assignment without propensity score weighting in order to assess the impact of ignoring potential confounding.

Data was simulated in *R* and was comprised of 15 binary latent class indicators, defining a 3-class structure, 8 normally distributed covariates, and a single normally-distributed distal outcome. The covariates, representing potential confounders, were associated with both latent class and the outcome; the strength of these associations was controlled through the α and β vectors of parameters, respectively. Specifically, the covariates for individual i in class c was generated as $\boldsymbol{X_{i,c}} = (\alpha_{1,c}X_{1i}, \alpha_{2,c}X_{2i}, \alpha_{3,c}X_{3i}, \alpha_{4,c}X_{4i}, \alpha_{5,c}X_{5i}, \alpha_{6,c}X_{6i}, \alpha_{7,c}X_{7i}, \alpha_{8,c}X_{8i})$ where X_{ni} represents a standard normal random variable (mean = 0, standard deviation = 1). The outcome for an individual i in class c was generated as $Y_{i,c} = \beta_{0,c} + \beta_{1,c}X_{1i,c} + \beta_{2,c}X_{2i,c} + \beta_{3,c}X_{3i,c} + \beta_{4,c}X_{4i,c} + \beta_{5,c}X_{5i,,c} + \beta_{6,c}X_{6i,c} + \beta_{7,c}X_{7i,c} + \beta_{8,c}X_{8i,c}$. Simulations investigated the effect of covarying class separation (i.e., entropy) and degree of confounding. Within a given class, all indicators had the same item probability (conceptually, "low," "medium," or "high"); the more distinct these item probabilities were across classes, the greater the class separation. We considered the following sets of item probabilities for Class 1, Class 2 and Class 3: (5%, 50%, 95%), (10%, 50%, 90%), (20%, 50%, 80%), and (30%, 50%, 70%). By varying the magnitude of both the α parameters (i.e., the coefficients linking the covariates and latent class) and β parameters (i.e., the coefficients linking the covariates and the distal outcome), we could control the magnitude of the confounding. For simplicity, within a given class, the α coefficients for all covariates and the β coefficients for all covariates were equal. We considered the following values for the α and β vectors, where larger values of α and β indicate greater confounding: ($\boldsymbol{\alpha}_1$=1, $\boldsymbol{\alpha}_2$=1, $\boldsymbol{\alpha}_3$=1; $\boldsymbol{\beta}_1$=1, $\boldsymbol{\beta}_2$=1, $\boldsymbol{\beta}_3$=1); ($\boldsymbol{\alpha}_1$=1, $\boldsymbol{\alpha}_2$=1.10, $\boldsymbol{\alpha}_3$=1.20; $\boldsymbol{\beta}_1$=1, $\boldsymbol{\beta}_2$=1.10, $\boldsymbol{\beta}_3$=1.20); ($\boldsymbol{\alpha}_1$=1, $\boldsymbol{\alpha}_2$=1.25, $\boldsymbol{\alpha}_3$=1.50; $\boldsymbol{\beta}_1$=1, $\boldsymbol{\beta}_2$=1.25, $\boldsymbol{\beta}_3$=1.50); and ($\boldsymbol{\alpha}_1$=1, $\boldsymbol{\alpha}_2$=1.50, $\boldsymbol{\alpha}_3$=2; $\boldsymbol{\beta}_1$=1, $\boldsymbol{\beta}_2$=1.50, $\boldsymbol{\beta}_3$=2). Each simulated dataset contained 5,000 observations and 1,000 simulations were performed at each setting.

The 1-step method we considered, Latent Class Causal Analysis (Kang & Schafer, 2010; Schafer & Kang, 2013) fits a joint model to the latent class indicators, potential confounders, and the distal outcome; we implemented this method using the *lcca* function in the LCCA package for *R,* specifying a 3-class model. In the *lcca* function, the user separately specifies covariates to control for with respect to the indicators and with respect to the outcome; we allowed all 8 covariates to predict both the indicators and the outcome. We implemented modal and pseudoclass assignment based on 3-class LCA results obtained using the *lca* function in the LCCA package. We obtained estimates of the Average Treatment Effect from the *lcca* package. Propensity scores, modeling class membership as based on modal or pseudoclass assignment, were estimated using logistic regression; propensity score weighting for multiple groups was conducted using the method described by McCaffrey et al. (2013) which fits *k* binary propensity score models for *k* treatment groups. In brief, we fit 3 binary propensity score models and for each individual calculated an inverse probability of treatment weight (IPTW); weights were trimmed at the 98th percentile to avoid extreme weights (Cole and Hernan, 2008). Inverse probability of treatment weighting also generates estimates of the Average Treatment Effect, making these results directly comparable to the results from LCCA. Differences in outcomes across classes were then estimated using propensity score weighted models that regressed the distal outcome on modal or pseudoclass assignment; this was implemented using the *survey* package in R (Lumley, 2004; Lumley, 2013). The pseudoclass method was performed 20 times for each dataset; final estimates were obtained using multiple imputation combining rules (Wang et al., 2005; Rubin, 1987). Unadjusted models were estimated by implementing the *lcca* function specifying no covariates and by implementing modal assignment without propensity score weighting. For the purposes of this simulation, all outcome and propensity score models were correctly specified.

Our primary interest was estimation of the three pairwise differences in outcome means across classes. For each simulation condition, we report bias, standard error (SE), mean squared error (MSE), and the 95% confidence interval coverage rate (i.e., the percentage of 95% confidence intervals that contained the true difference in means) averaged across the three pairwise differences.

Results

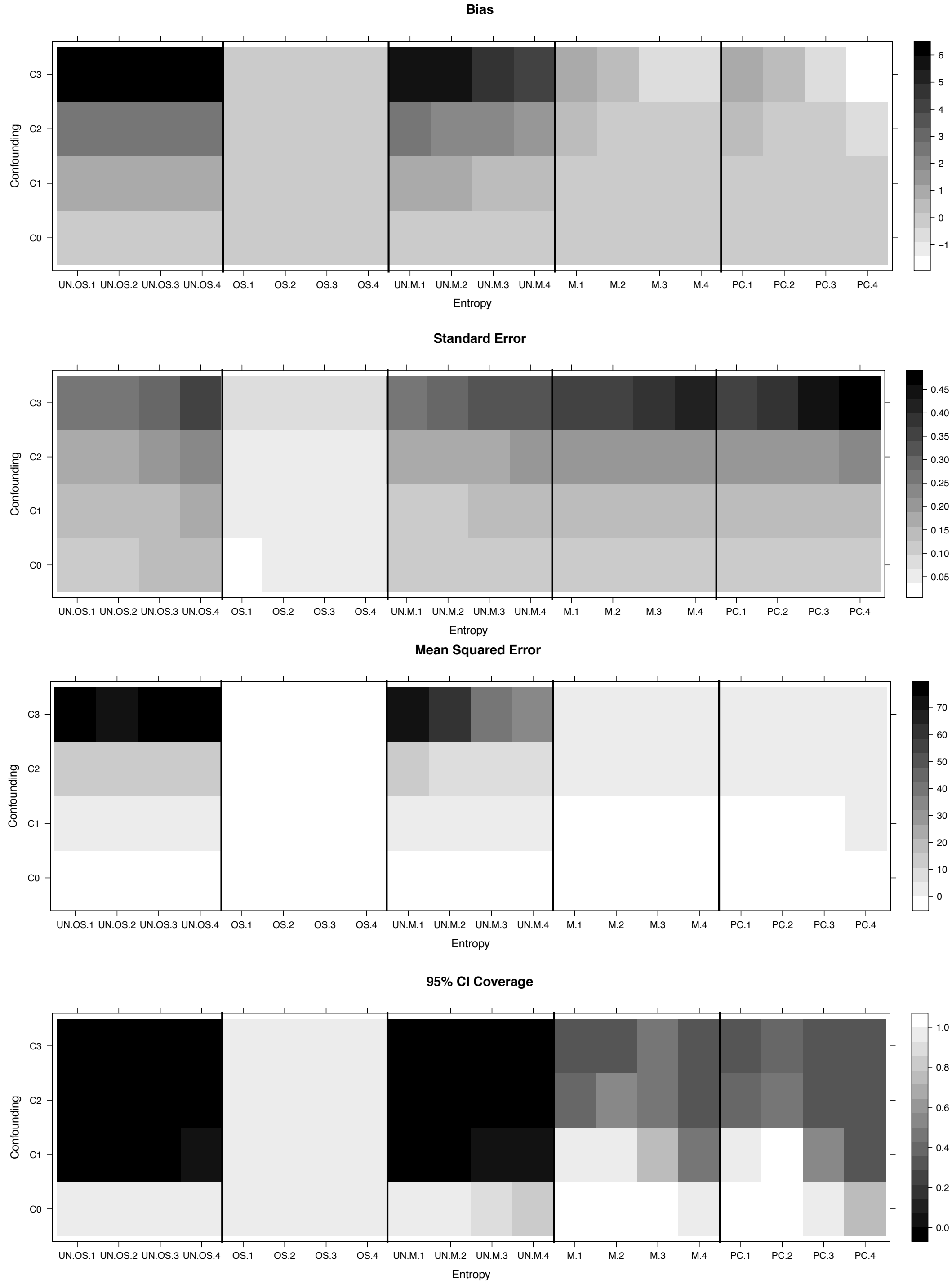
Figure 4.4 presents 4 figures depicting bias, standard error, mean squared error, and 95% confidence interval coverage rates for each method as a function of both entropy and degree of confounding (numerical results presented in Table 4.1 and Table 4.2). With regard to bias, both unadjusted methods yielded bias that was orders of magnitude larger than any of the three adjusted methods. The bias for unadjusted LCCA was primarily affected by the degree of confounding, whereas the bias for unadjusted modal assignment increased as both confounding increased and class separation worsened (i.e., entropy decreased). Adjusted LCCA showed very small bias regardless of the degree of confounding or entropy. Modal and pseudoclass assignment with propensity score weighting both showed notable bias reductions compared to the unadjusted methods, yet the bias for these methods was consistently larger than for adjusted LCCA and increased with confounding and worsening class separation.

With regard to standard error, adjusted LCCA consistently yields the smallest SE estimates, while the other four methods yield SE estimates that are approximately 3-4 times larger in magnitude. When there is no confounding (denoted C0 in Figure 4.4) or minimal confounding (denoted C1), these four methods have similar SEs; as confounding increases (C2 and C3 in our simulations), both 3-step methods with propensity score weighting yield notably larger SEs. For all methods, SEs increase as entropy decreases; the magnitude of this increase is smallest for adjusted LCCA.

Mean squared error estimates are very small for adjusted LCCA, and show gradual increases as confounding increases and entropy decreases. MSE estimates for unadjusted methods are extremely large in comparison, primarily reflecting large bias arising from these methods. Both 3-step methods with propensity score weighting yielded significantly smaller MSE estimates compared to the unadjusted methods. MSE estimates for both modal and pseudoclass assignment with propensity score weighting yielded MSE estimates that were significantly smaller than the unadjusted methods, yet were still orders of magnitude larger than adjusted LCCA, particularly when the degree of confounding was large. MSE for the 3-step methods was driven both by notable bias and larger SE estimates.

The final panel in Figure 4.4 presents 95% confidence interval coverage estimates. In general, the adjusted LCCA method yields conservative coverage (greater than 97% for all conditions) and is not significantly affected by degree of confounding or entropy. The unadjusted methods show close to nominal coverage under no confounding and high entropy (C0E1, C0E2), yet coverage for these methods is 0% under almost all conditions that involve confounding (C1, C2, and C3). Coverage rates for the 3-step methods with propensity score weighting are also conservative (near 100%) when there is little or no confounding and high entropy; however, coverage notably decreases as entropy decreases and confounding increases. Both of these methods show quite poor coverage rates under the two conditions with the most confounding.

Figure 4.4. Bias, standard error, mean squared error and 95% confidence interval coverage as a function of both entropy and degree of confounding



Abbreviations: UN.OS= Unadjusted one-step (LCCA); OS = One-step (LCCA); UN.M = Unadjusted Modal assignment; M = Modal assignment + PS weighting; PC = pseudoclass assignment + PS weighting. 1, 2, 3, and 4 suffixes denote average entropy of approximately 0.50, 0.70, 0.90 and 0.96, respectively. C0, C1, C2, C3 denote confounding of (α_1=1, α_2=1, α_3=1; β_1=1, β_2=1, β_3=1); (α_1=1, α_2=1.10, α_3=1.20; β_1=1, β_2=1.10, β_3=1.20); (α_1=1, α_2=1.25, α_3=1.50; β_1=1, β_2=1.25, β_3=1.50); (α_1=1, α_2=1.50, α_3=2; β_1=1, β_2=1.50, β_3=2). In all figures dark shading indicates worse performance.

Table 4.1 Average Bias and Standard Error across the three pairwise contrasts across classes.

	Average Bias				*Average SE*			
	C0E1	**C0E2**	**C0E3**	**C0E4**	**C0E1**	**C0E2**	**C0E3**	**C0E4**
1-step	0.000	0.000	0.000	0.000	0.035	0.037	0.043	0.054
Unadj 1-step	0.001	-0.002	-0.002	-0.001	0.106	0.110	0.128	0.152
Unadj Modal	-0.007	-0.022	-0.060	-0.079	0.104	0.104	0.104	0.104
Modal + PSW	-0.008	-0.020	-0.059	-0.078	0.104	0.104	0.105	0.104
Pseudo + PSW	-0.009	-0.024	-0.075	-0.113	0.104	0.104	0.105	0.105
	C1E1	**C1E2**	**C1E3**	**C1E4**	**C1E1**	**C1E2**	**C1E3**	**C1E4**
1-step	0.001	0.000	0.000	0.002	0.037	0.038	0.043	0.050
Unadj 1-step	0.878	0.879	0.877	0.882	0.127	0.132	0.154	0.186
Unadj Modal	0.839	0.786	0.611	0.530	0.126	0.127	0.129	0.130
Modal + PSW	0.028	-0.010	-0.128	-0.180	0.129	0.130	0.132	0.133
Pseudo + PSW	0.027	-0.021	-0.173	-0.284	0.129	0.131	0.134	0.136
	C2E1	**C2E2**	**C2E3**	**C2E4**	**C2E1**	**C2E2**	**C2E3**	**C2E4**
1-step	0.001	0.001	0.006	0.011	0.044	0.044	0.046	0.048
Unadj 1-step	2.493	2.492	2.495	2.494	0.169	0.174	0.201	0.238
Unadj Modal	2.400	2.266	1.845	1.640	0.171	0.174	0.183	0.189
Modal + PSW	0.219	0.112	-0.210	-0.341	0.190	0.194	0.202	0.207
Pseudo + PSW	0.221	0.084	-0.323	-0.607	0.192	0.198	0.211	0.218
	C3E1	**C3E2**	**C3E3**	**C3E4**	**C3E1**	**C3E2**	**C3E3**	**C3E4**
1-step	0.001	0.004	0.007	0.005	0.070	0.071	0.071	0.072
Unadj 1-step	5.980	5.972	5.981	5.981	0.263	0.270	0.306	0.349
Unadj Modal	5.770	5.457	4.508	4.046	0.275	0.286	0.314	0.330
Modal + PSW	1.036	0.645	-0.376	-0.729	0.341	0.362	0.394	0.405
Pseudo + PSW	1.046	0.555	-0.670	-1.438	0.352	0.390	0.438	0.461

Abbreviations: C0, C1, C2, C3 denote confounding of (α_1=1, α_2=1, α_3=1; β_1=1, β_2=1, β_3=1); (α_1=1, α_2=1.10, α_3=1.20; β_1=1, β_2=1.10, β_3=1.20); (α_1=1, α_2=1.25, α_3=1.50; β_1=1, β_2=1.25, β_3=1.50); (α_1=1, α_2=1.50, α_3=2; β_1=1, β_2=1.50, β_3=2). E1, E2, E3, and E4 suffixes denote average entropy of approximately 0.50, 0.70, 0.90 and 0.96, respectively.

Table 4.2 Average Mean Squared Error and 95% Confidence Interval coverage rates across the three pairwise contrasts across classes.

	Average MSE				*Average Coverage*			
	C0E1	**C0E2**	**C0E3**	**C0E4**	**C0E1**	**C0E2**	**C0E3**	**C0E4**
1-step	0.001	0.001	0.001	0.002	98.4%	98.0%	97.6%	97.5%
Unadj 1-step	0.011	0.012	0.015	0.021	95.1%	95.4%	95.5%	95.7%
Unadj Modal	0.011	0.012	0.018	0.023	95.0%	93.8%	87.1%	82.1%
Modal + PSW	0.001	0.002	0.008	0.014	100.0%	100.0%	99.9%	97.4%
Pseudo + PSW	0.001	0.002	0.012	0.027	100.0%	100.0%	99.2%	74.2%
	C1E1	**C1E2**	**C1E3**	**C1E4**	**C1E1**	**C1E2**	**C1E3**	**C1E4**
1-step	0.001	0.001	0.001	0.002	98.0%	97.8%	97.5%	96.6%
Unadj 1-step	1.560	1.567	1.566	1.596	0.0%	0.0%	0.1%	0.7%
Unadj Modal	1.427	1.255	0.769	0.584	0.0%	0.0%	0.9%	2.6%
Modal + PSW	0.013	0.012	0.043	0.076	99.2%	99.9%	76.7%	48.7%
Pseudo + PSW	0.013	0.012	0.069	0.171	99.0%	100.0%	50.3%	30.8%
	C2E1	**C2E2**	**C2E3**	**C2E4**	**C2E1**	**C2E2**	**C2E3**	**C2E4**
1-step	0.001	0.001	0.001	0.002	98.9%	98.6%	98.3%	97.8%
Unadj 1-step	12.627	12.615	12.651	12.653	0.0%	0.0%	0.0%	0.0%
Unadj Modal	11.708	10.440	6.951	5.482	0.0%	0.0%	0.0%	0.0%
Modal + PSW	0.297	0.213	0.239	0.386	41.2%	51.0%	43.8%	33.7%
Pseudo + PSW	0.304	0.202	0.356	0.865	41.6%	48.3%	33.7%	32.2%
	C3E1	**C3E2**	**C3E3**	**C3E4**	**C3E1**	**C3E2**	**C3E3**	**C3E4**
1-step	0.002	0.002	0.025	0.069	99.4%	99.7%	99.5%	99.2%
Unadj 1-step	74.241	74.047	74.289	74.306	0.0%	0.0%	0.0%	0.0%
Unadj Modal	69.158	61.806	42.161	33.727	0.0%	0.0%	0.0%	0.0%
Modal + PSW	3.065	1.688	0.919	1.726	33.1%	35.4%	44.6%	34.2%
Pseudo + PSW	3.138	1.442	1.512	4.701	33.1%	38.9%	35.5%	29.4%

Abbreviations: C0, C1, C2, C3 denote confounding of (α_1=1, α_2=1, α_3=1; β_1=1, β_2=1, β_3=1); (α_1=1, α_2=1.10, α_3=1.20; β_1=1, β_2=1.10, β_3=1.20); (α_1=1, α_2=1.25, α_3=1.50; β_1=1, β_2=1.25, β_3=1.50); (α_1=1, α_2=1.50, α_3=2; β_1=1, β_2=1.50, β_3=2). E1, E2, E3, and E4 suffixes denote average entropy of approximately 0.50, 0.70, 0.90 and 0.96, respectively.

4.9 Part II: Motivating Example

Methods

In Part II of this study, we applied the five previously discussed methods to our substantive question of interest regarding the effect of classes of substance use treatment services on substance use problems. This analysis used data on 5,527 youth ages 12-18 who were receiving exclusively outpatient services through treatment providers funded through the Substance Abuse and Mental Health Administration's Center for Substance Abuse Treatment's (CSAT) discretionary grant funding program. Youth were enrolled in one of 9 CSAT-funded treatment programs: the Effective Adolescent Treatment (EAT) program that supported MET/CBT-5 implementation (Dennis et al., 2004; Melchior et al., 2007; SAMHSA, 2003); the Cannabis Youth Treatment experiment (Dennis et al., 2004; Diamond et al., 2002) which randomized youth to MET, CBT, family-based therapy, or MDFT, or Adolescent Community Reinforcement Approach (ACRA); the Adolescent Treatment Models program (Dennis et al., 2003) providing community-based care; the Adolescent Residential Treatment (SAMHSA, 2002) providing continuing care for discharged youth; the Strengthening Communities' Youth program (Dennis et al., 2008) aimed at building partnerships among community, school-based and juvenile justice treatment services for early intervention, outpatient and intensive outpatient programs; the Targeted Capacity Expansion program (Wilson et al., 2005) providing intensive outpatient and inpatient services; the Young Offenders Reentry Program (SAMHSA, 2004) providing services to youth re-entering the community; the Family and Juvenile Treatment Drug Court program (SAMHSA, 2005) providing comprehensive services through drug courts; and the Assertive Adolescent and Family Treatment program (Godley et al., 2007) promoting family-centered services. For study participation, parents provided written informed consent and adolescents provided assent; institutional review boards approved the study protocol at each site.

Youth were assessed with the Global Appraisal of Individual Needs (GAIN; Dennis, 1999), a comprehensive survey that assesses the following domains: demographics, substance use and substance use treatment, risk behaviors, mental and physical health, legal status, environment risk factors, and education/vocation status. All data collected with the GAIN are based on youth self-report; reliability studies conducted by Dennis et al. (1999, 2000, 2002) have found very good reliability statistics for the majority of the GAIN indices (i.e., Cronbach's α greater than 0.85). The GAIN's Treatment Received Scale (TxRS) was used to assess the substance use treatment services that youth received from study enrollment to 3 months; this 20-item scale includes subscales that measure provision of Direct (i.e., individual-focused), Family, and External (i.e., case management) services (Dennis et al., 2010). A total of 12 items, 4 from each of the subscales, were used as latent class indicators. The distal outcome of interest is the change in Substance Problems Scale (SPS) score from baseline to 3 months.

Adjusted analysis controlled for the potential confounders including demographic variables (age, sex, and race/ethnicity (White, Black, Hispanic, and Other)); baseline substance use variables (prior substance use treatment, recognition of substance problems, Substance Frequency Scale (past 90 days), Substance Problems Scale (past month), Substance Dependence Scale (past year), and Treatment Motivation Index); legal status variables (any justice system involvement, any arrests, any days in a controlled environment (all with respect to past 90 days), and the Crime Violence Scale); and mental health variables (days affected by emotional problems (past 90 days), and the Behavioral Complexity Scale).

Analyses for adjusted LCCA, modal assignment with propensity score weighting, and pseudoclass assignment with propensity score weighting, included the 12 latent class indicators, the distal outcome (SPS change score), and the 15 potential confounders. Analyses using the unadjusted LCCA model and modal assignment included only the 12 latent class indicators and the distal outcome. A 4-class model was specified for all methods, based on the results of our previous latent class analysis (see Chapter 2). The same covariates were included in the LCCA model as were included in the propensity score models. We present all six of the estimated pairwise differences in distal outcomes between classes.

Results

As Table 4.3 shows, unadjusted and adjusted estimates vary significantly with regard to the resulting statistical inference. The unadjusted 1-step method suggests that the Individual-Focused Services class, the Individual- and Family-Focused Services class, and the Multiple Services class each have significantly larger decreases on the Substance Problems Scale from baseline to 3 months than the Low Service Utilization class (respective estimates are -0.37, p-val=0.04; -0.60, p-val=0.001; and -0.59, p-val=0.01). Similarly, the unadjusted analysis based on modal assignment also suggests that the Individual- and Family-Focused Services class and the Multiple Services class each have significantly larger decreases on the SPS than the Low Service Utilization class (respective estimates are -0.49, p-val=0.004; and -0.45, p-val=0.03). However, none of the adjusted methods (LCCA, modal, or pseudoclass assignment) show any significant differences across classes with regard to changes in SPS.

This example highlights that conducting latent class analysis with distal outcomes with and without controlling for potential confounding can lead to notably different substantive interpretations. Specifically, the unadjusted analyses suggest that youth do better when receiving fewer types of substance use treatment services and that adding additional services does not provide any additional benefits. However, our adjusted analyses suggest that these results merely reflect the fact that youth who received the fewest number of services types actually had fewer substance problems at baseline, and thus the fact that they received fewer service types may be a reflection of their lower treatment need (i.e., it may be due to confounding). When adjusting for baseline characteristics, we find that there is no significant difference across the groups, indicating that perhaps the natural process of treatment self-selection or referral is effectively resulting in youth receiving equally effective treatment services, given their baseline need.

Table 4.3. Estimated class differences in the 3-month Substance Problems Scale, as estimated by three methods that adjust for potential confounding and two unadjusted methods

	1-Step, Unadjusted			*1-Step, Covariate Adjusted*		
Class Comparison	*Est*	*SE*	*p-val*	*Est*	*SE*	*p-val*
Indiv v Low	-0.370	0.182	0.042*	-0.267	0.345	0.440
(Indiv & Fam) v Low	-0.603	0.180	0.001*	-0.372	0.288	0.196
Multiple v Low	-0.594	0.229	0.010*	-0.274	0.359	0.445
(Indiv & Fam) v Indiv	-0.233	0.143	0.102	-0.106	0.183	0.564
Multiple v Indiv	-0.224	0.190	0.239	-0.008	0.258	0.977
Multiple v (Indiv & Fam)	0.009	0.210	0.965	0.098	0.281	0.727

	Modal Assignment, Unadjusted			*Modal Assignment, Propensity Score Weighted*			*Pseudoclass Assignment, Propensity Score Weighted*		
Class Comparison	*Est*	*SE*	*p-val*	*Est*	*SE*	*p-val*	*Est*	*SE*	*p-val*
Indiv v Low	-0.303	0.165	0.066	-0.102	0.183	0.578	-0.116	0.184	0.530
(Indiv & Fam) v Low	-0.486	0.167	0.004*	-0.200	0.183	0.274	-0.198	0.181	0.273
Multiple v Low	-0.447	0.207	0.031*	-0.196	0.238	0.410	-0.236	0.240	0.324
(Indiv & Fam) v Indiv	-0.183	0.108	0.089	-0.098	0.112	0.380	-0.082	0.115	0.473
Multiple v Indiv	-0.144	0.163	0.377	-0.094	0.189	0.618	-0.121	0.188	0.522
Multiple v (Indiv & Fam)	0.039	0.166	0.812	0.004	0.188	0.984	0.082	0.115	0.473

Note: * denotes p-values < 0.05
Abbreviations: Low = Low Service Utilization class; Indiv = Individual-Focused Services class; Indiv & Fam = Individual- and Family-Focused Services class; Multiple = Multiple Services class

4.10 Discussion

Overall, the results from both our simulation study and our motivating example of adolescents in substance use treatment demonstrate that effect estimates from latent class analysis with distal outcomes may vary substantially whether or not potential confounding is adjusted for. Confounding in settings where all variables are fully observed is widely recognized and addressed statistically, yet recognition of and statistical methods for confounding in latent variable regression are only recently emerging. Controlling for confounding in latent variable regression presents unique challenges, particularly when the latent variable is the treatment of interest. In this paper we examine a recently proposed 1-step method, LCCA, which addresses confounding through joint modeling of the latent class indicators, confounders, and the distal outcome. Additionally, we examine methods to incorporate propensity score weighting with classical 3-step methods, namely modal and pseudoclass assignment.

In general, our results indicate that LCCA performs quite well under a wide range of conditions, yielding very small bias, reasonable standard errors, and small MSE estimates. Coverage rates were somewhat conservative in our simulation results. However, implementation of LCCA, or other 1-step methods, may not be feasible in all settings. In our work, we frequently encountered convergence issues. Additionally, in some cases, the latent class estimation under a 1-step method may be unduly influence by the distal outcome. Initially, we considered an additional substance use outcome, the Substance Frequency Scale. However, implementation of LCCA with this outcome yielded a notably different 4 class structure, with regard both to conditional item probabilities and estimated class prevalences, than our original 4-class model. Note that these classes were well preserved both for the latent class regression model discussed in our previous work (see Chapter 2) and also when modeling the Substance Problems Scale as a distal outcome. We were unable to present results regarding the Substance Frequency Scale given that the 1-step and 3-step results were not comparable with our other results due to differences in the estimated latent class structure. Conceptually, it is undesirable for the distal outcome to significantly influence the latent classes, particularly when the goal is to estimate the causal effect of class membership on the distal outcome; Petras and Masyn (2010) further discuss this limitation of 1-step methods for distal outcomes.

Although we found that modal and pseudoclass assignment with propensity score weighting were able to significantly reduce the bias in the effect estimates in the presence of confounding, these approaches do have some limitations. They yielded biased estimates, due to misclassification of the predicted class with respect to the true class; this bias was small for conditions that involved less confounding and good class separation (high entropy). However, our simulation results showed that these methods perform poorly when there are high levels of confounding and, as has been shown previously in the case of latent variable regression, when entropy decreases, due to the increased rate of misclassification of individuals (Asparouhov & Muthén, 2013; Bolck et al., 2004; Vermunt, 2010). Another important consideration is that the propensity score, like the final effect estimate, is calculated with respect to the predicted, rather than the true, latent class. Thus, propensity score weighting balances *predicted* latent classes on baseline covariates; given misclassification, balance on predicted latent classes does not necessarily imply balance on true latent classes. By adding an additional estimation step that relies on predicted latent class, these methods do introduce additional bias. This can be seen by comparing the simulation results for the condition with no confounding (C0) to the conditions with confounding – for a given entropy level, these methods show larger bias in the presence of confounding relative to no confounding. This indicates that the strategy of adjusting for confounding using propensity scores based on predicted class does introduce additional bias beyond the bias arising from the use of predicted class in the outcome model. Although we feel it is important to assess covariate balance with respect to the predicted class, and only proceed with the final analysis if classes are sufficiently well balanced, it is clear that balancing predicted classes does not necessarily balance true classes. However, this limitation is inherent to the nature of latent variables: since latent classes are unobserved, we can never estimate the propensity score or assess balance with regard to the true latent class.

Given that the 3-step methods do significantly reduce confounding, future work will explore methods to reduce the bias and improve coverage rates of these 3-step approaches. As we discuss in Chapter 3, numerous methods have been proposed in recent years to correct the bias from 3-step methods in the context of standard latent variable regression – see Bakk et al. (2013), Vermunt (2010), Asparouhov & Muthén (2013); Petersen et al. (2013), and Lanza et al. (2013) for details on these correction methods. In the absence of confounding, these corrected 3-step methods perform quite similarly to 1-step methods with respect to bias, SE, MSE and 95% confidence interval converge, yet are often less computationally intensive. Potentially these correction methods could be used in conjunction with propensity score methods. Again, in the case of distal outcomes, 3-step methods are particularly advantageous since they allow estimation of the latent classes without influence from the distal outcome. Another potential method for addressing confounding in this context was proposed by Yamaguchi (n.d.) and is also based on propensity score weighting, but obtains effect estimates by calculating potential outcomes for each treatment class through weighting and then taking pairwise differences in potential outcomes across classes. For simplicity and given that corrected 3-step methods have not yet been widely adopted by applied researchers, we chose to focus only on classical 3-step methods in this study. However, future work should investigate the potential for combining corrected 3-step methods with propensity score methods, as well as Yamaguchi's proposed method.

Although our simulation results highlight the performance of the 1-step method, we do think that a 3-step approach ultimately offers greater modeling flexibility. One limitation to our simulation study was that for all conditions and all methods, both the latent class / covariate model and the covariate / distal outcome model were correctly specified. It is possible that propensity score based 3-step approach would be more advantageous relative to a 1-step model in the case of model misspecification. Given the parametric constraints of the joint model specified in the LCCA package, it is likely in practice that one or both of the association models will be incorrectly specified. Propensity score methods have been shown to be relatively robust to model misspecification relative to covariate adjustment; additional strategies to buffer the effects of model misspecification include non-parametric estimation of the propensity score (Lee et al., 2009; McCaffrey et al., 2004; Stuart, 2010) and using doubly robust estimation (Kang & Schafer, 2007). Thus, one significant advantage of a 3-step approach is that it allows the incorporation of propensity score methods, which may perform better under some conditions, particularly model misspecification, than the covariate adjustment implemented in 1-step methods.

Overall, this paper highlights that applied researchers should think critically about confounding in the context of latent variable regression; as in contexts with fully observed variables, failure to adjust for potential confounders may lead to significantly biased results and potentially misleading inferences. Although methodological development in this area has been limited so far, given the complications of latent treatment groups, we discuss three proposed methods, a 1-step approach as well as 3-step approaches that include propensity score weighting. As we discuss, each of these approaches do reduce confounding bias, the 1-step method more effectively than the 3-step methods, yet each approach has limitations. Future methodological work should focus on developing and refining methods that can address confounding for latent class analysis with distal outcomes, and assess performance under a broader array of conditions, including model misspecification.

CHAPTER 5. DISCUSSION

The substantive objective that motivated this dissertation work was identifying adolescent drug treatment services that effectively improve substance use outcomes *in practice*. Our approach to answering this question was unique in two regards: first, we defined our treatment groups with respect to services that youth *actually reported receiving*, and second, we used latent class analysis to empirically identify subgroups of youth who were receiving similar types of treatment services. Most existing studies regarding the effectiveness of adolescent drug treatment have defined treatment groups in terms of programs or services that youth were *randomized t*o or *enrolled in*. Given the notable rates of attrition and noncompliance among adolescents in treatment, as well as factors such as poor program fidelity or service provision through multiple entities, program enrollment may not accurately reflect the services that youth actually receive. For this work, our treatment groups definitions were based on 12 self-reported items from the Global Appraisal of Individual Needs (GAIN) survey that ask whether youth received specific drug treatment services; this approach ensured that we classified youth based on the services they reported receiving, rather than program enrollment. Additionally, empirically identifying treatment groups through latent class analysis revealed that youth enrolled in different treatment programs reported receiving the same class of treatment services, indicating that a latent class approach is an important strategy for data reduction when analyzing databases with youth enrolled in numerous treatment programs.

This dissertation also addressed the methodological challenges associated with latent variable regression. Latent variable regression with predictive auxiliary variables is one of the most popular applications, although recent methodological work has increasingly focused on latent variable regression with a distal outcome. Since distal outcome methods are still relatively new, there are many unanswered methodological questions. Broadly, latent variable regression methods can be described as 1-step or 3-step methods. Although 1-step methods are typically described as the optimal approach for predictive auxiliary variables, this is less true for distal outcomes. One important concern is that the use of a joint model, as in 1-step methods, allows the distal outcome to influence latent class formation. When estimating the causal effect of latent classes on a distal outcome, it is not desirable for the distal outcome, which presumably occurred subsequent to when the latent class indicators were measured, to influence class formation. Furthermore, when there are several distal outcomes of interest, if class formation is notably influenced by the distal outcome, a 1-step method becomes intractable; classes, and thus effect estimates, are not comparable across distal outcomes. We encountered this challenge in Chapter 4 when assessing two outcomes of interest, the Substance Frequency Scale (SFS) and the Substance Problems Scale (SPS); the 1-step method (with 4 classes) with the SFS identified markedly different classes from all of our other analyses, and thus we did not present these results. Given these limitations of 1-step methods, as well as the potential for convergence issues due to the complexity of the joint model likelihood, 3-step methods may be preferable in some contexts. Classical 3-step methods have been shown to be biased for many latent class regression applications; our findings regarding classical 3-step methods in Chapter 3 are consistent with previous findings. More recently, corrected 3-step methods have been proposed; our findings in Chapter 3 suggest these methods perform quite well, and similarly to 1-step methods, with regard to bias, mean squared error, and 95% confidence interval coverage rates.

An additional methodological challenge in latent variable regression that is only recently receiving attention is the problem of confounding. In latent variable regression with a distal outcome, the latent class often represents the treatment or intervention of interest. These analyses are frequently performed using observational data, and thus there is the possibility that an individual's baseline characteristics may be associated with his or her latent class membership. Thus, adjusting for potential confounders is essential when estimating the causal effect of the latent class on the distal outcome. Although methods to address confounding when all variables are fully observed, such as propensity score methods, are well developed, only recently have methods been proposed for addressing confounding when the treatment is latent. In Chapter 4 we examine the performance of three approaches – a 1-step method, as well as two different 3-step methods combined with propensity score weighting.

5.1 Summary of main findings

Aim 1. This study included 5,527 adolescents who were receiving outpatient drug treatment services throughout the US through 9 different treatment programs funded by the Substance Abuse and Mental Health Service Administration's Center for Substance Abuse Treatment. Using 12 items from the GAIN's Treatment Received Scale that reflected individual-focused, family-based, and case management services, we performed latent class analysis to identify classes of treatment services youth reported receiving. Four latent classes were identified: Class 1: Low Service Utilization (12% of youth); Class 2: Individual-Focused Services (39%); Class 3: Individual- and Family-Focused Services (38%); and Class 4: Multiple Services (11%). Latent class regression identified significant predictors of latent class membership, including demographics, factors related to substance use, and justice system involvement; for example, youth in Class 1 had the lowest levels of baseline substance use, and youth in Class 4 had higher levels of baseline substance use as well as greater involvement with the justice system. Additionally, we found that each of the 4 treatment classes was estimated to contain youth from all 9 of the treatment programs, reflecting both the heterogeneity in service provision within a given treatment program, as well as the similarity of services provided across treatment programs.

Aim 2. This study reviewed the conceptual differences between 1-step, classical 3-step, and recent corrected 3-step methods for latent variable regression, particularly in the case of distal outcomes. Using a simulation study, we compared the statistical performance of 5 methods, namely a 1-step method, modal assignment, multiple pseudoclass assignment, posterior probability regression, and Vermunt's correction with modal assignment. The estimates of interest were the pairwise differences in outcome means across classes; statistical performance was assessed in terms of bias, standard error, mean squared error, and 95% confidence interval coverage. We found that the 1-step method and the corrected 3-step both performed quite well across all conditions, but modal and pseudoclass assignment yielded significantly biased estimates, attenuated standard error estimates, and poor coverage rates, particularly when entropy was low. Overall, although 1-step methods perform well under many conditions, they may not always be conceptually appropriate, especially one is interested in causal inference with respect to distal outcomes. Therefore, 3-step methods may often be more appropriate or computationally tractable; in these cases, corrected 3-step methods should be used.

Aim 3. This study examined a recent 1-step method for addressing confounding when regressing a distal outcome on latent classes, Kang and Schafer's (2010) Latent Class Causal Analysis (LCCA). We also explored the use of propensity score weighting in combination with both modal and pseudoclass assignment. We compared the statistical performance of these three methods to adjust for confounding, as well as both an unadjusted 1-step and 3-step method, on simulated data. We found that the 1-step method, modal assignment with propensity score weighting, and pseudoclass assignment with propensity score weighting all significantly reduced the bias due to confounding relative to the two unadjusted methods. The 1-step method performed better with regard to bias and 95% confidence interval coverage, especially in the case of low entropy (i.e., poor class separation). Additionally, we applied these methods to our adolescent drug treatment data in order to estimate the effects of drug treatment class on substance use outcomes. We found that both unadjusted methods indicated significant differences across treatment classes with respect to the Substance Problems Scale, yet no significant differences were found using any of the three adjusted methods, suggesting that the unadjusted results may be attributable to confounding. Overall, potential confounding should be carefully considered when conducting latent variable regression with a distal outcome, and future work should continue to develop statistical methods to address confounding in this context.

5.2 Synthesis of findings

Overall, this study offers an interesting finding with regard to adolescent substance use treatment. Our results suggest that once baseline differences among youth who are receiving difference classes of treatment services are controlled for, youth have similar outcomes on the Substance Problems Scale. This suggests that, perhaps contrary to some conventional wisdom, that increasing the number of treatment services types that youth receive does not necessarily improve outcomes. The fact that all treatment groups showed a decline in SPS, yet groups did not differ significantly, suggests that providers may be doing an effective job in practice of matching youth with appropriate treatment services. Future work should expand on these findings by investigating other substance use outcomes, outcomes at subsequent follow-up visits, and different samples of youth.

Additionally, this study adds to the literature regarding 1-step and 3-step methods for latent variable regression, and our results in Chapter 3 highlight the potential for corrected 3-step methods for latent variable regression with distal outcomes. Additionally, we conduct one of the first comparisons of statistical methods to address confounding in the context of latent variable regression with distal outcomes. We investigated the potential of combining propensity score weighting with classical 3-step methods, and found that the use of propensity scores did indeed significantly reduce the bias arising from confounding. However, the magnitude of the bias for these methods was still quite large relative to the results from the comparative 1-step method. Thus, the results of Chapters 3 and 4 together suggest the potential for future methodological work to combine propensity score methods and corrected 3-step methods.

5.3 Strengths and limitations of the findings

Strengths. A notable strength of this work was our use of observational data on of a large number of adolescents receiving outpatient drug treatment services from providers throughout the US. Although this data is not truly nationally representative, it offers a chance to examine the effectiveness of treatment services provided by typical community-based providers, rather than the more tightly controlled environments of randomized studies.

Another strength of this study was the use of simulation studies to compare the statistical performance of various analysis methods. The simulation study presented in Chapter 3 represents one of the more comprehensive simulation studies conducted to date with regard to latent variable regression on a distal outcome, and includes the use of Vermunt's correction as just recently implemented in Mplus. Furthermore, previous simulation studies have primarily investigated the effects of entropy and sample size; we additionally considered the effects of the magnitude of class differences and the magnitude of variance in the distal outcome. In Chapter 4 we conducted one of the only simulation studies assessing the relative performance of methods to address confounding in the context of latent variable regression with a distal outcome. The simulation study in Chapter 4 is also supplemented with an applied example using our adolescent substance use data; this provides a concrete example that further highlights the problems of confounding in this context for applied researchers.

Limitations. One limitation for the analyses involving the adolescent substance use data was the fact that all data, including the indicators used for latent class formation as well as substance use outcomes, were self-reported. The veracity of self-reported data may be affected by recall bias, social desirability bias, or other factors. Inaccuracies in these data may have resulted in misclassification of youth with regard to treatment classes or with regard to substance use outcomes; inaccuracies may be more of a concern for outcomes, which may be perceived by youth as more sensitive information. Unfortunately, data that could be used to verify youth self-reports, such as provider reports or administrative records of service provision, were not available. Furthermore, we lacked information regarding the frequency of services received, quality metrics of services, and the authority providing the services; such data could enrich the descriptions of our latent classes. Finally, the descriptions of treatment classes in our study may not be fully generalizable to other populations of adolescents receiving outpatient treatment services.

A limitation of our simulation studies, like all simulation studies, is that their results may not generalize to all applications of the methods in practice. Although we attempted to generate data that were reflective of our adolescent substance use data and data commonly encountered by social and behavioral researchers, our simulated data by necessity represents a simplified case. For example, we generated data with a 3 class structure, whereas our applied example had a 4 class structure; it is highly plausible, yet not guaranteed, that our results from the 3 class case will generalize to the 4 class case. Furthermore, we only considered conditions in which all aspects of our statistical models were correctly specified with regard to how the data was generated; our simulation results may not be applicable in the context of model misspecification.

5.4 Future directions

Given that statistical methods for addressing confounding when the treatment of interest is a latent variable have only recently been proposed, there are many possible directions for future research. As we alluded to previously, future work should continue to investigate incorporating propensity score methods with existing methods for latent variable regression with distal outcomes. Particularly, recent corrected 3-step, when used in conjunction with propensity score methods, may offer a very flexible and powerful means of addressing confounding in this context. Propensity scores methods offer many superior statistical properties relative to covariate adjustment, which 1-step methods in this context rely on. Given the breadth of propensity score methodology, future work should also investigate alternate propensity score methods, such as matching or subclassification, or techniques such as non-parametric estimation of the propensity scores or doubly robust estimation of the effect estimates. Finally, a very important consideration for future work is to examine the robustness to model misspecification of both 1-step methods and 3-step methods combined with propensity score methods.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716-723.

Akers, R. L., Krohn, M. D., Lanza-Kaduce, L., Radosevich, M. (1979). Social learning and deviant behavior: A specific test of a general theory. *American Sociological Review, 44,* 636-655.

Asparouhov, T., & Muthén, B. O. (2013). Auxiliary variables in mixture modeling: A 3-step approach using Mplus (Mplus Web Notes: No. 15, Version 7). http://www.statmodel.com/examples/webnotes/webnote15.pdf.

Bakk, Z., Tekle, F., & Vermunt, J. K. (2013). Estimating the association between latent class membership and external variables using bias adjusted three-step approaches. *Sociological Methodology, 43,* 272-311.

Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz. P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association, 92,* 1375–1386.

Bolck, A., Croon, M. A., & Hagenaars, J. A. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political Analysis, 12,* 3-27.

Bollen, K. A. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika, 61,* 109-121.

Butera, N. M., Lanza, S. T., & Coffman, D. L. (2013). A framework for estimating causal effects in latent class analysis: Is there a causal link between early sex and subsequent profiles of delinquency? *Prevention Science* [epub ahead of print].

Clark, S., & Muthén, B. O. (2009). Relating latent class analysis results to variables not included in the analysis. https://www.statmodel.com/download/relatinglca.pdf

Clogg, C. C. (1995). Latent class models: Recent developments and prospects for the future. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp.311-352). New York, NY: Plenum.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. New York, NY: Wiley.

Conduct Problems Prevention Research Group, (2013). Trajectories of risk for early sexual activity and early substance use in the fast track prevention program. *Prevention Science*, [Epub ahead of print].

Croon, M. (2002). Using predicted latent scores in general latent structure models. In G. A. Marcoulides, & I. Moustaki (Eds.), *Latent Variable and Latent Structure Models* (pp. 195–223). Mahwah, NJ: Erlbaum Associates.

Dennis, M. L. (1999). *Global Appraisal of Individual Needs (GAIN) Administration guide for the GAIN and related measures (Version 1299)*. Bloomington, IL: Chestnut Health Systems.

Dennis, M. L., Dawud-Noursi, S., Muck, R. D., & McDermeit, M. (2003). The need for developing and evaluating adolescent treatment models. In S. J. Stevens, & A. R. Morral (Eds.), *Adolescent substance abuse treatment in the United States: Exemplary models from a national evaluation study* (pp. 3–34). Binghamton, NY: The Haworth Press.

Dennis, M., Godley, S. H., Diamond, G., Tims, F. M., Babor, T., Donaldson, J.,..., & Funk, R. (2004). The Cannabis Youth Treatment (CYT) study: Main findings from two randomized trials. *Journal of Substance Abuse Treatment, 27,* 197-213.

Dennis, M. L., Ives, M., Funk, R., Modisette, K., Bledsaw, R., & Ihnes, P. (2010). *GAIN-I encyclopedia of supplemental documentation on scales and other calculated variables.* Bloomington, IL: Chestnut Health Systems.

Dennis, M. L., Ives, M. L., White, M. K., & Muck, R. D. (2008). The Strengthening Communities for Youth (SCY) initiative: A cluster analysis of the services received, their correlates and how they are associated with outcomes. *Journal of Psychoactive Drugs,* *40,* 3-16.

Dennis, M.L., Scott, C.K., Godley, M.D., & Funk, R. (1999). Comparisons of adolescents and adults by ASAM profile using GAIN data from the Drug Outcome Monitoring Study (DOMS): Preliminary data tables, Vol. 2010. Bloomington, IL: Chestnut Health Systems.

Dennis, M.L., Scott, C.K., Godley, M.D., & Funk, R.R. (2000). Predicting outcomes in adult and adolescent treatment with case mix vs. level of care: Findings from the Drug Outcome Monitoring Study. *Drug and Alcohol Dependence, 60(Suppl. 1)*, s51.

Dennis, M. L., Titus, J. C., Diamond, G., Donaldson, J., Godley, S. H., Tims, F.M.,..., & The C. Y. T. Steering committee. (2002). The Cannabis Youth Treatment (CYT) experiment: rationale, study design, and analysis plans. *Addiction, 97,* 16-34.

Diamond, G., Godley, S. H., Liddle, H. A., Sampl, S., Webb, C., Tims, F. M., & Meyers, R. (2002). Five outpatient treatment models for adolescent marijuana use: A description of the Cannabis Youth Treatment Interventions. *Addiction, 97*, 70–83.

Diamond, G., & Josephson, A. (2005). Family-based treatment research: A 10-year update. *Journal of American Academy of Child and Adolescent Psychiatry, 44(9),* 872-887.

Donders, A. R., van der Heijden, G. J., Stijnen, T., & Moons, K. G. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology, 59,* 1087-1091.

Feingold, A., Tiberio, S. S., & Capaldi, D.M. (2013). New approaches for examining associations with latent categorical variables: Applications to substance abuse and aggression. *Psychology of Addictive Behaviors*, [epub ahead of print].

Godley, M. D., & Godley, S. H. (2011). Assertive Continuing Care for Adolescents. In J. F. Kelly, & W. L. White (Eds.), *Addiction recovery management: Theory, research and* *practice*, (pp.103-126). New York, NY: Humana Press.

Godley, M. D., Godley, S. H., Dennis, M. L., Funk, R., & Passetti, L. L. (2002). Preliminary outcomes from the assertive continuing care experiment for adolescents discharged from residential treatment. *Journal of Substance Abuse Treatment*, *23(1),* 21-32.

Godley, M. D., Godley, S. H., Dennis, M. L., Funk, R. R., & Passetti, L. L. (2007). The effect of assertive continuing care on continuing care linkage, adherence and abstinence following residential treatment for adolescents with substance use disorders. *Addiction, 102,* 81-93.

Godley, S. H., Godley, M. D., Pratt, A., & Wallace, J. L. (1994). Case management services for substance abusers: A program description. *Journal of Substance Abuse Treatment, 11(4),* 309-317.

Godley, M. D., Kahn, J. H., Dennis, M. L., Godley, S. H., & Funk, R. (2005). The stability and impact of environmental factors on substance use and problems after adolescent outpatient treatment for cannabis abuse and dependence. *Psychology of Addictive Behaviors, 19(1),* 62-70.

Goodman, L.A. (2007). On the assignment of individuals to latent classes. *Sociological Methodology, 37,* 1–22.

Gudicha, D. W., & Vermunt, J. K. (in press). Mixture model clustering with covariates using adjusted three-step approaches. In B. Lausen, D. van den Poel, & A. Ultsch (Eds.), *Algorithms from and for nature and life: Studies in classification, data analysis, and knowledge organization*. Heidelberg, Germany: Springer-Verlag.

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis*. Cambridge: Cambridge University Press.

Hallfors, D., Cho, H., Sanchez, V., Khatapoush, S., Kim, H. M., & Bauer, D. (2006). Efficacy vs effectiveness trial results of an indicated “model” substance abuse program: implications for public health. *American Journal of Public Health, 96*, 2254-2259.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (1998). *Multisystemic Treatment of Antisocial Behavior in Children and Adolescents*. New York, NY: The Guilford Press.

Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P.B. (2009). *Multisystemic Therapy for Antisocial Behavior in Children and Adolescents* (2nd ed.). New York, NY: The Guilford Press.

Higgins, S. T., Silverman, K., & Heil, S. H. (2008). *Contingency management in substance abuse treatment*. New York, NY: The Guilford Press.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15,* 199-236.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika, 87(3)*, 706-710.

Johnston, L. D., O'Malley, P. M., Bachman, J. G., & Schulenberg, J. E. (2013). Monitoring the Future national results on drug use: 2012 Overview, Key Findings on Adolescent Drug Use. Ann Arbor: Institute for Social Research, The University of Michigan.

Kang, J., & Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science, 22,* 523–539.

Kang, J., & Schafer, J. L. (2010). Estimating average treatment effects when the treatment is a latent class. Technical Report 10-05. Retrieved from http://stat.psu.edu/research-old/technical-reports/2010-technical-reports/TR10-05.pdf/view.

Kelly, J. F., & Myers, M.G., (2007). Adolescents' participation in Alcoholics Anonymous and Narcotics Anonymous: review, implications and future directions. *Journal of*

Psychoactive Drugs, 39, 259–269.

Landa, R. J., Gross, A. L., Stuart, E. A., & Bauman, M. (2012). Latent class analysis of early developmental trajectory in baby siblings of children with autism. *Journal of Child Psychology and Psychiatry, 53(9),* 986-996.

Lanza, S. T., Coffman, D. L., Xu, S. (2013). Causal inference in latent class analysis. *Structural Equation Modeling, 20,* 361–383.

Lanza, S. T., Tan, X., & Bray, B. C. (2013). Latent class analysis with distal outcomes: A flexible model-based approach. *Structural Equation Modeling, 20,* 1-26.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.

Lee, B. K., Lessler, J., & Stuart, E. A. (2009). Improving propensity score weighting using machine learning. *Statistics in Medicine, 29(3),* 337-346.

Liddle, H. A. (1999). Theory development in a family-based therapy for adolescent drug abuse. *Journal of Clinical Child Psychology, 28,* 521-532.

Liddle, H. A. (2004). Family-based therapies for adolescent alcohol and drug use research contributions and future research needs. *Addiction, 99,* 76-92.

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association, 90,* 1112-1121.

Lu, I. R. R., & Roland, T. D. (2008). Avoiding and correcting bias in score-based latent variable regression with discrete manifest items. *Structural Equation Modeling, 15,* 462-490.

Lumley. T. (2004). Analysis of complex survey samples. *Journal of Statistical Software, 9(1)*, 1-19.

Lumley, T. (2013). Package 'survey': R package version 3.29-5. Retrieved from http://cran.r-project.org/web/packages/survey/survey.pdf.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F.

(2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine, 32(19),* 3388-3414.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods, 9,* 403-425.

Melchior, L. A., Griffith, A. A., & Huba, G. J. (2007). Evaluation of the Effective Adolescent Treatment (EAT) program. Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Substance Abuse Treatment.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behaviors*. New York, NY: The Guilford Press.

Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14,* 535-569.

Oberski, D. L. & Satorra, A. (2013). Measurement error models with uncertainty about the error variance. *Structural Equation Modeling, 20,* 409-428.

Petersen, J., Bandeen-Roche, K., Budtz-Jørgensen, E., & Groes Larsen, K. (2012). Predicting latent class scores for subsequent analysis. *Psychometrika, 77,* 244–262.

Petras, H., & Masyn, K. (2010). General growth mixture analysis with antecedents and consequences of change. In A. R. Piquero, & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 69-100). New York, NY: Springer.

Petry, N. M., & Simic, F. (2002). Recent advances in the dissemination of contingency management techniques: clinical and research perspectives. *Journal of Substance Abuse Treatment, 23,* 81-86.

Riley, K. J., Srikanth, P., Choi, D., & McCarty, D. (2012). Treatment length and outcomes among adolescents: a secondary data analysis. *Substance Abuse Treatment, Prevention, and Policy, 7,* 1-12.

Rowe, C. L., & Liddle, H. A. (2003). Substance abuse. *Journal of Marital and Family Therapy, 29,* 97-120.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: J. Wiley & Sons.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services & Outcomes Research Methodology, 2,* 169-188.

Sánchez, B. N., Budtz-Jørgensen, E., & Ryan, L. M. (2009). An estimating equations approach to fitting latent exposure models with longitudinal health outcomes. *Annals of Applied Statistics*, *3*, 830–856.

Schafer, J. L., & Kang, J. (2013). LCCA package for R users' guide (Version 1.1.0). University Park: The Methodology Center, Penn State. Retrieved from http://methodology.psu.edu/downloads/lcca.

Schwartz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*, 333-343.

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *4*, 563–576.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward.

Statistical Science, 25(1), 1-21.

Substance Abuse and Mental Health Services Administration, Office of Applied Studies. (2012). Results from the 2011 National Survey on Drug Use and Health: Summary of national findings (DHHS Publication No. SMA 12-4713). Rockville, MD: Office of Applied Studies, Substance Abuse and Mental Health Services Administration.

Substance Abuse and Mental Health Services Administration, 2002. *Center for Substance Abuse Treatment: Grants to Improve the Quality and Availability of Residential Treatment and its Continuing Care Component for Adolescents.* Request for Applications (RFA) no. TI-02-007. Part I-Programmatic guidance. Retrieved from http://www.samhsa.gov/grants/content/2002/ti02007_adoles.html.

Substance Abuse and Mental Health Services Administration, 2003. *Center for Substance Abuse Treatment: Adopt/expand effective adolescent alcohol and drug abuse treatment.* Request for Applications (RFA) no. TI03-007. Part I-Programmatic guidance. Retrieved from http://www.samhsa.gov/grants/content/2003/ti03007_eat.htm.

Substance Abuse and Mental Health Services Administration, 2004. *Center for Substance Abuse Treatment: Substance Abuse Treatment and Reentry Services to Sentenced Juveniles and Young Adult Offenders Returning to the Community from the Correctional System.* Request for Applications (RFA) no. TI04-001. Part I-Programmatic guidance. Retrieved from http://www.samhsa.gov/grants/2004/nofa/ti04002rfa_yorp.htm.

Substance Abuse and Mental Health Services Administration, 2005. *Center for Substance Abuse Treatment: Family and Juvenile Treatment Drug Courts.* Request for Applications (RFA) no. TI05-005. Part I-Programmatic guidance. Retrieved from http://www.samhsa.gov/grants/2005/ti_05_005.aspx.

Vanderplasschen, W., Rapp, R. C., Wolf, J. R., & Broekaert, E. (2004). The development and implementation of case management for substance use disorders in North America and Europe. *Psychiatric Services, 55(8),* 913-922.

Vermunt, J. K., (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis, 18*, 450–469.

Wang, C. P., Brown, C. H., & Bandeen-Roche, K. (2005). Residual diagnostics for growth mixture models: Examining the impact of a preventive intervention on multiple trajectories of aggressive behavior. *Journal of the American Statistical Association, 100(471),* 1054–1076.

Wang, C. Y., Huang, Y., Chao, E. C., & Jeffcoat, M. K. (2008) Expected estimating equations for missing data, measurement error, and misclassification, with application to longitudinal nonignorable missing data. *Biometrics, 64,* 85-95.

Wilson, M. T., Atanda, R., Atkinson, D. D., & Mulvey, K. (2005). Outcomes from the targeted capacity expansion (TCE) substance abuse treatment program. *Evaluation and Program Planning, 28,* 341-348.

Winters, K. C., Botzet, A. M., & Fahnhorst, T. (2011). Advances in adolescent substance abuse treatment. *Currents Psychiatry Reports, 13,* 416-421.

Yamaguchi, K. (n.d.). Causal analysis for linear regression with a latent-class treatment variable: An application to an analysis of IT equipment use on income. Retrieved from http://popcenter.uchicago.edu/links/LAT_TREAT_NEW7r.pdf

APPENDIX I. R Code for Aim 2 Simulations

```
##############################################################
# LC --> distal outcome
# Multiple replications compare 1-step and 3-step methods
# Reads in data generated by Mplus (.dat files)
# 1-step performed in LCCA package
# 3-step: modal + propensity scores
# 3-step: pseudoclass + propensity scores
# Estimates Treatment effects (Class Differences)
# Returns Std Bias, SE, Std MSE, 95% Coverage rates
# Updated: Aug 1, 2013
##############################################################
#  libraries
library(lcca)
library(survey)
library(mitools)
##############################################################

run.sim <- function(num.sim=1000, my.N=5000, my.mu, my.K=20, sim.id){

  results <- NULL

  for(i in 1:num.sim)
  {
    data<-read.table(paste(sim.id,"data",i,".dat",sep=""))
    colnames(data) <-c("U1","U2","U3","U4","U5","U6","U7","U8","U9",
       "U10","U11","U12","U13","U14","U15", "Y_obs","modal")

    ### Convert 0=No --> 2=No for LCCA package
    data$U1[data$U1==0] <- 2
    data$U2[data$U2==0] <- 2
    data$U3[data$U3==0] <- 2
    data$U4[data$U4==0] <- 2
    data$U5[data$U5==0] <- 2
    data$U6[data$U6==0] <- 2
    data$U7[data$U7==0] <- 2
    data$U8[data$U8==0] <- 2
    data$U9[data$U9==0] <- 2
    data$U10[data$U10==0] <- 2
    data$U11[data$U11==0] <- 2
    data$U12[data$U12==0] <- 2
    data$U13[data$U13==0] <- 2
    data$U14[data$U14==0] <- 2
    data$U15[data$U15==0] <- 2

    o1<-onestep.lcr(data=data, mu=my.mu)
    o2<-threestep.lcr(data=data, mu=my.mu, K=my.K)

    results.new<-c(o1,o2)
    results<-rbind(results,results.new)
  }

  colnames(results)<-c("os.2v1.b","os.2v1.SE","os.2v1.MSE","os.2v1.cov",
                       "os.3v1.b","os.3v1.SE","os.3v1.MSE","os.3v1.cov",
                       "os.3v2.b","os.3v2.SE","os.3v2.MSE","os.3v2.cov",
                       "p1","p2","p3",
                       "m.2v1.b","m.2v1.SE","m.2v1.MSE","m.2v1.cov",
                       "m.3v1.b","m.3v1.SE","m.3v1.MSE","m.3v1.cov",
                       "m.3v2.b","m.3v2.SE","m.3v2.MSE","m.3v2.cov",
                       "p.2v1.b","p.2v1.SE","p.2v1.MSE","p.2v1.cov",
                       "p.3v1.b","p.3v1.SE","p.3v1.MSE","p.3v1.cov",
                       "p.3v2.b","p.3v2.SE","p.3v2.MSE","p.3v2.cov",
```

```
                     "p1","p2","p3")
  return(results)

}

################################################################
##  functions!!
################################################################

################################################################
################################################################
## 1-step code
################################################################
################################################################
onestep.lcr <- function(data, mu, num.tx=3, num.item=15, iter=100000){

  ### Calculate true treatment effects
  diff.21 <- mu[2]-mu[1]
  diff.31 <- mu[3]-mu[1]
  diff.32 <- mu[3]-mu[2]
  ### FIT THE LCCA model
  fit.lcca <- lcca(formula.treatment =
       cbind(U1,U2,U3,U4,U5,U6,U7,U8,U9,U10,U11,U12,U13,U14,U15)~ 1,
        formula.outcome = Y_obs~1, data=data,nclass=num.tx,iter.max=iter,
        flatten.rhos=1,stabilize.alphas=1,se.method="STANDARD")
  ### Resolve class switching
  cond.mean <- NULL
  # Sum conditional probs (rhos) within each class to determine class names
  cond.mean[1]<-sum(fit.lcca$theta[1:num.item])
  cond.mean[2]<-sum(fit.lcca$theta[(num.item+1):(2*num.item)])
  cond.mean[3]<-sum(fit.lcca$theta[(2*num.item+1):(3*num.item)])
  # Check if classes are well differentiated
  test <- pairwise.diffs(cond.mean)
  if(sum(abs(test)<0.25)>0) print("Warning: Indistinguishable Classes")
  #Reorder classes and summarize lcca
  ord <- order(cond.mean)
  fit.lcca.ord <- permute.class(fit.lcca, ord)
  ### Save Average Treatment Effects, Bias, SE, and MSE
  output <- NULL
  os.2v1<-summary(fit.lcca.ord)[94][[1]][1]
  os.2v1.b<-(summary(fit.lcca.ord)[94][[1]][1]-diff.21)/diff.21
  os.2v1.SE<-summary(fit.lcca.ord)[94][[1]][4]
  os.2v1.MSE <- (diff.21*os.2v1.b)^2
  os.2v1.cov<-(os.2v1-1.96*os.2v1.SE)<diff.21 & diff.21<(os.2v1+1.96*os.2v1.SE)

  os.3v1<-summary(fit.lcca.ord)[94][[1]][2]
  os.3v1.b<-(summary(fit.lcca.ord)[94][[1]][2]-diff.31)/diff.31
  os.3v1.SE<-summary(fit.lcca.ord)[94][[1]][5]
  os.3v1.MSE <- (diff.31*os.3v1.b)^2
  os.3v1.cov<-(os.3v1-1.96*os.3v1.SE)<diff.31 & diff.31<(os.3v1+1.96*os.3v1.SE)

  os.3v2<-summary(fit.lcca.ord)[94][[1]][3]
  os.3v2.b<-(summary(fit.lcca.ord)[94][[1]][3]-diff.32)/diff.32
  os.3v2.SE <-summary(fit.lcca.ord)[94][[1]][6]
  os.3v2.MSE <- (diff.32*os.3v2.b)^2
  os.3v2.cov<-(os.3v2-1.96*os.3v2.SE)<diff.32 & diff.32<(os.3v2+1.96*os.3v2.SE)
  # Save Class Prevalences
  p1<-fit.lcca.ord$marg.gamma[1]
  p2<-fit.lcca.ord$marg.gamma[2]
  p3<-fit.lcca.ord$marg.gamma[3]

  output<-as.numeric(c(os.2v1.b,os.2v1.SE,os.2v1.MSE,os.2v1.cov,
                       os.3v1.b,os.3v1.SE,os.3v1.MSE,os.3v1.cov,
```

```
                       os.3v2.b,os.3v2.SE,os.3v2.MSE,os.3v2.cov,
                       p1,p2,p3))
  return(output)

}
################################################################
################################################################
## 3-step code
################################################################
################################################################
threestep.lcr<-function(data, mu, num.cov=0, num.item=15, num.tx=3, K=20,
iter=100000){

  ### define N = number of individuals
  N <- dim(data)[1]
  ### Calculate true treatment effects
  diff.21 <- mu[2]-mu[1]
  diff.31 <- mu[3]-mu[1]
  diff.32 <- mu[3]-mu[2]

  ### ESTIMATE LATENT CLASSES WITH LCA
  fit.lca <- lca(cbind(U1,U2,U3,U4,U5,U6,U7,U8,U9,U10,U11,U12,U13,U14,U15)~1,
             nclass=num.tx,data=data,iter.max=iter,flatten.rho=1,
             flatten.gammas=1)
  ### Resolve class switching
  cond.mean <- NULL
  # Sum conditional probs (rhos) within each class to determine class names
  cond.mean[1]<-sum(fit.lca$theta[1:num.item])
  cond.mean[2]<-sum(fit.lca$theta[(num.item+1):(2*num.item)])
  cond.mean[3]<-sum(fit.lca$theta[(2*num.item+1):(3*num.item)])
  # Check if classes are well differentiated
  test <- pairwise.diffs(cond.mean)
  if(sum(abs(test)<0.25)>0) print("Warning: Indistinguishable Classes")
  #Reorder classes and summarize lca
  ord <- order(cond.mean)
  fit.lca.ord <- permute.class(fit.lca, ord)
  ### SAVE POSTERIOR PROBABILITIES
  probs <- fit.lca.ord$post.probs
  data <- cbind(data, probs)  ### "Class 1", "Class 2", "Class 3"
  ################################################################
  ## CLASS ASSIGNMENT via MODAL ASSIGNMENT CLASS
  ################################################################
  data$modal_c[data$"Class 1">data$"Class 2"&data$"Class 1">data$"Class 3"]<-1
  data$modal_c[data$"Class 2">data$"Class 1"&data$"Class 2">data$"Class 3"]<-2
  data$modal_c[data$"Class 3">data$"Class 1"&data$"Class 3">data$"Class 2"]<-3
  ################################################################
  ## CLASS ASSIGNMENT via PSEUDOCLASS DRAWS
  ################################################################
  # matrix for pseudoclass draws
  obs.tx <- matrix(nrow=N, ncol=K)
  prob <- cbind(data$"Class 1", data$"Class 2", data$"Class 3")

  for(i in 1:N)
  {   # set up multinomial RV and take K draws
      # rmultinom generates matrix (dim = num.tx x K)
      T <- rmultinom(K, size = 1, prob=prob[i,])
      # Convert draws matrix to vector, Save observed tx class
      for(k in 1:K)
          {    obs.tx[i,k] <- which.max(T[,k])
          }
  }
  colnames(obs.tx) <- c("pc1","pc2","pc3","pc4","pc5","pc6","pc7",
       "pc8","pc9","pc10","pc11","pc12","pc13","pc14","pc15","pc16",
```

```
      "pc17","pc18","pc19","pc20")
  data <- cbind(data,obs.tx)
  ################################################################
  ## ESTIMATE TX EFFECTS: Modal
  ################################################################
  # outcome regression for Y ~ Tx (Tx is factor variable)
  data$modal_c <- factor(data$modal_c, levels=c(1,2,3))
  out1m <- lm(Y_obs ~ C(modal_c,contr.treatment(3,base=1)), data=data)
  out2m <- lm(Y_obs ~ C(modal_c,contr.treatment(3,base=2)), data=data)
  ## Class 2 v Class 1
  m.2v1 <- summary(out1m)$coefficient[2,1]
  m.2v1.b <- (m.2v1-diff.21)/diff.21
  m.2v1.SE <- summary(out1m)$coefficient[2,2]
  m.2v1.MSE <- (diff.21*m.2v1.b)^2
  m.2v1.cov <- (m.2v1-1.96*m.2v1.SE)<diff.21 & diff.21<(m.2v1+1.96*m.2v1.SE)
  ## Class 3 v Class 1
  m.3v1 <- summary(out1m)$coefficient[3,1]
  m.3v1.b <- (m.3v1-diff.31)/diff.31
  m.3v1.SE <- summary(out1m)$coefficient[3,2]
  m.3v1.MSE <- (diff.31*m.3v1.b)^2
  m.3v1.cov <- (m.3v1-1.96*m.3v1.SE)<diff.31 & diff.31<(m.3v1+1.96*m.3v1.SE)
  ## Class 3 v Class 2
  m.3v2 <- summary(out2m)$coefficient[3,1]
  m.3v2.b <- (m.3v2-diff.32)/diff.32
  m.3v2.SE <- summary(out2m)$coefficient[3,2]
  m.3v2.MSE <- (diff.32*m.3v2.b)^2
  m.3v2.cov <- (m.3v2-1.96*m.3v2.SE)<diff.32 & diff.32<(m.3v2+1.96*m.3v2.SE)
  ################################################################
  ## ESTIMATE TX EFFECTS: Pseudoclass
  ################################################################
  # Matrix for estimated treatment effects
  q.matrix <- matrix(nrow=K, ncol=choose(num.tx,2))
  se.matrix <- matrix(nrow=K, ncol= choose(num.tx,2))
  pc_names <- c("pc1","pc2","pc3","pc4","pc5","pc6","pc7","pc8","pc9","pc10",

"pc11","pc12","pc13","pc14","pc15","pc16","pc17","pc18","pc19","pc20")
  ### For each pseudoclass draw, estimate Tx Effect and SE
  for(k in 1:K)
  {  # Define temporary variable: kth pseudoclass Tx Class
     data$tx_temp <- factor(data[,pc_names[k]],levels = c(1,2,3))
     #  outcome regression for Y ~ Tx
     out1p <- lm(Y_obs ~ C(tx_temp,contr.treatment(3,base=1)), data=data)
     out2p <- lm(Y_obs ~ C(tx_temp,contr.treatment(3,base=2)), data=data)
     ## Class 2 v Class 1 contrast, SE
     q.matrix[k,1] <- summary(out1p)$coefficient[2,1]
     se.matrix[k,1] <- summary(out1p)$coefficient[2,2]
     ## Class 3 v Class 1 contrast, SE
     q.matrix[k,2] <- summary(out1p)$coefficient[3,1]
     se.matrix[k,2] <- summary(out1p)$coefficient[3,2]
     ## Class 3 v Class 2 contrast, SE
     q.matrix[k,3] <- summary(out2p)$coefficient[3,1]
     se.matrix[k,3] <- summary(out2p)$coefficient[3,2]
  }
  ### Combine Tx Effects, SE across pseudoclasses. Then calc Bias, MSE, 95% CI
coverage
  output <- NULL
  p.2v1 <-
as.numeric(MIcombine(as.list(q.matrix[,1]),as.list(se.matrix[,1]))[1])
  p.2v1.b <- (p.2v1 - diff.21)/diff.21
  p.2v1.SE <-
as.numeric(MIcombine(as.list(q.matrix[,1]),as.list(se.matrix[,1]))[2])
  p.2v1.MSE <- (diff.21*p.2v1.b)^2
  p.2v1.cov <- (p.2v1-1.96*p.2v1.SE)<diff.21 & diff.21<(p.2v1+1.96*p.2v1.SE)
```

```
  p.3v1<-as.numeric(MIcombine(as.list(q.matrix[,2]),as.list(se.matrix[,2]))[1])
  p.3v1.b <- (p.3v1 - diff.31)/diff.31
  p.3v1.SE<-
as.numeric(MIcombine(as.list(q.matrix[,2]),as.list(se.matrix[,2]))[2])
  p.3v1.MSE <- (diff.31*p.3v1.b)^2
  p.3v1.cov <- (p.3v1-1.96*p.3v1.SE)<diff.31 & diff.31<(p.3v1+1.96*p.3v1.SE)

  p.3v2<-
as.numeric(MIcombine(as.list(q.matrix[,3]),as.list(se.matrix[,3]))[1])
  p.3v2.b <- (p.3v2 - diff.32)/diff.32
  p.3v2.SE<-
as.numeric(MIcombine(as.list(q.matrix[,3]),as.list(se.matrix[,3]))[2])
  p.3v2.MSE <- (diff.32*p.3v2.b)^2
  p.3v2.cov <- (p.3v2-1.96*p.3v2.SE)<diff.32 & diff.32<(p.3v2+1.96*p.3v2.SE)
  ### Class prevalences from LCA (same for Modal and Pseudoclass)
  p1<-summary(fit.lca.ord)[66][[1]][1]
  p2<-summary(fit.lca.ord)[66][[1]][2]
  p3<-summary(fit.lca.ord)[66][[1]][3]

  output<-as.numeric(c(m.2v1.b,m.2v1.SE,m.2v1.MSE,m.2v1.cov,
                       m.3v1.b,m.3v1.SE,m.3v1.MSE,m.3v1.cov,
                       m.3v2.b,m.3v2.SE,m.3v2.MSE,m.3v2.cov,
                       p.2v1.b,p.2v1.SE,p.2v1.MSE,p.2v1.cov,
                       p.3v1.b,p.3v1.SE,p.3v1.MSE,p.3v1.cov,
                       p.3v2.b,p.3v2.SE,p.3v2.MSE,p.3v2.cov,
                       p1,p2,p3))
  return(output)

}
###############################################################
###############################################################
## PAIRWISE Diff function
###############################################################
###############################################################
## CALCULATE PAIRWISE DIFFS
## Solve class switching by comparing sum of estimated item prob
## across classes. Since data is set up with monotonically increasing
## item probs across classes, sum(C1 item) < sum(C2 items) < sum(C3 items)
## This function compares pairwise differences of these sums to
## ensure that they are sufficiently large: i.e. classes are
## adequately differentiated
###############################################################

pairwise.diffs <- function(x)
{   # create matrix of combination pairs
    prs <- cbind(rep(1:length(x), each = length(x)), 1:length(x))
    # drop ones that compare same classes
    drops <- NULL
    for(i in 1:(length(x))^2 )
       { if (prs[i,1]==prs[i,2]) drops <- c(drops, i)
       }
    new <- prs[-drops,]
    # do pairwise differences
    result <- x[new[,1]] - x[new[,2], drop = FALSE]
}
###############################################################
```

APPENDIX II. R Code for Aim 3 Simulations

```
##############################################################
# Xs---> LC --> distal outcome
# Multiple replications compare 1-step and 3-step methods
# Generates and then analyzes datasets
# 1-step performed in LCCA package
# Unadj 3-step: modal, no covariate adjustment
# 3-step: modal + propensity scores
# 3-step: pseudoclass + propensity scores
# Estimates Treatment effects (Class Differences)
##############################################################
#  libraries
library(lcca)
library(survey)
library(mitools)
##############################################################

aim3.sim <- function(num.sim=1000, my.N=5000, my.tx, my.mu, my.p, my.noise,
                     my.x.mean, my.x.sd, my.alpha1, my.alpha2, my.alpha3,
                     my.beta1, my.beta2, my.beta3, my.K=20, sim.id){

  results <- NULL

  for(i in 1:num.sim)
  {
    data<-a3.data(N=my.N, tx.prob=my.tx, mu=my.mu, p=my.p, noise=my.noise,
                  x.mean=my.x.mean, x.sd=my.x.sd,
                  alpha1=my.alpha1, alpha2=my.alpha2, alpha3=my.alpha3,
                  beta1=my.beta1, beta2=my.beta2, beta3=my.beta3)

    write.table(data,paste(sim.id,"A3dataset",i,".txt",sep=""),row.names=F)

    o1<-onestep.lcca(data=data)
    o2<-threestep.lcca(data=data, K=my.K, p=my.p)

    results.new<-c(o1,o2)
    results<-rbind(results,results.new)

  }

  return(results)

}

##############################################################
##  functions!!
##############################################################
### Generate simulated data for AIM 3###
### 3 latent classes, equal proportions
### 8 covariates: 4 continuous, 4 binary
### Continuous potential outcomes Y1, Y2, Y3 for each class
### Can vary:  (1) class differentiation (item probabilities)
###            (2) strength of Class ~ X association (ALPHAS)
###            (3) strength of Y ~ X association (BETAS)

a3.data <- function(N,num.tx=3,num.item=15,num.cov=8,tx.prob=c(0.33,0.33,0.34),

x.mean,x.sd,alpha1,alpha2,alpha3,beta1,beta2,beta3,mu,p,noise=1)
{
  # initialize variables for TRUE treatment class
  tx1.ind <- tx2.ind <- tx3.ind <- tx <- rep(NA,N)
  # treatment group is determined by multinomial draw
```

```
T <- t(rmultinom(N, size=1, prob=tx.prob))
# rmultinom generates matrix: convert matrix T into vector tx
tx1.ind <-  T[,1]==1
tx2.ind <-  T[,2]==1
tx3.ind <-  T[,3]==1
tx[tx1.ind] <- 1
tx[tx2.ind] <- 2
tx[tx3.ind] <- 3
#####################
### Class and Xs related by ALPHAs
X1 <- rnorm(N, mean=x.mean[1], sd=x.sd[1])
X2 <- rnorm(N, mean=x.mean[2], sd=x.sd[2])
X3 <- rnorm(N, mean=x.mean[3], sd=x.sd[3])
X4 <- rnorm(N, mean=x.mean[4], sd=x.sd[4])
X5 <- rnorm(N, mean=x.mean[5], sd=x.sd[5])
X6 <- rnorm(N, mean=x.mean[6], sd=x.sd[6])
X7 <- rnorm(N, mean=x.mean[7], sd=x.sd[7])
X8 <- rnorm(N, mean=x.mean[8], sd=x.sd[8])
X.matrix <- cbind(X1,X2,X3,X4,X5,X6,X7,X8)
C1.matrix <- X.matrix * alpha1
C2.matrix <- X.matrix * alpha2
C3.matrix <- X.matrix * alpha3
X.obs <- matrix(nrow=N,ncol=num.cov)
X.obs[tx1.ind,] <- C1.matrix[tx1.ind,]
X.obs[tx2.ind,] <- C2.matrix[tx2.ind,]
X.obs[tx3.ind,] <- C3.matrix[tx3.ind,]
colnames(X.obs) <- c("X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8")
#####################
### generate continuous potential outcomes for each class: Y1, Y2, Y3
### Ys are function of Intercept (mu) and BETAs*Xs
Y1 <- mu[1] + apply(beta1*X.obs,1,sum) + rnorm(N,mean=0,sd=noise)
Y2 <- mu[2] + apply(beta2*X.obs,1,sum) + rnorm(N,mean=0,sd=noise)
Y3 <- mu[3] + apply(beta3*X.obs,1,sum) + rnorm(N,mean=0,sd=noise)
# Expected values of each potential outcome
E.Y1 <- mean(Y1)
E.Y2 <- mean(Y2)
E.Y3 <- mean(Y3)
# These are true treatment effects (create vectors of length N)
diff.21 <- rep(E.Y2-E.Y1, N)
diff.31 <- rep(E.Y3-E.Y1, N)
diff.32 <- rep(E.Y3-E.Y2, N)
# observed outcome is potential outcome under individual's true class
Y_obs <- NULL
Y_obs[tx1.ind] <- Y1[tx1.ind]
Y_obs[tx2.ind] <- Y2[tx2.ind]
Y_obs[tx3.ind] <- Y3[tx3.ind]
### generate latent class indicators
U.matrix <- matrix(nrow = N, ncol= num.item)
### Create "Potential" latent class indicators for C1, C2, C3
item1.matrix <- array(rbinom(n=N*num.item,1,prob=p[1]), dim=c(N,num.item))
item2.matrix <- array(rbinom(n=N*num.item,1,prob=p[2]), dim=c(N,num.item))
item3.matrix <- array(rbinom(n=N*num.item,1,prob=p[3]), dim=c(N,num.item))
### Convert 0 --> 2 for LCCA package
item1.matrix[item1.matrix==0] <- 2
item2.matrix[item2.matrix==0] <- 2
item3.matrix[item3.matrix==0] <- 2
### Observe indicators that correspond to true tx class
U.matrix[tx1.ind,] <- item1.matrix[tx1.ind,]
U.matrix[tx2.ind,] <- item2.matrix[tx2.ind,]
U.matrix[tx3.ind,] <- item3.matrix[tx3.ind,]
colnames(U.matrix) <-c("U1","U2","U3","U4","U5","U6","U7","U8","U9",
     "U10","U11","U12","U13","U14","U15")
```

```
  data <- data.frame(cbind(tx,Y1,Y2,Y3,Y_obs,diff.21,diff.31,
       diff.32,X.obs,U.matrix))
}
###############################################################
###############################################################
## 1-step code
###############################################################
###############################################################
onestep.lcca <- function(data, num.tx=3, num.item=15, iter=100000){

  ### Calculate true treatment effects
  diff.21 <- mean(data$diff.21)
  diff.31 <- mean(data$diff.31)
  diff.32 <- mean(data$diff.32)
  fit.unadj <-
lcca(formula.treatment=cbind(U1,U2,U3,U4,U5,U6,U7,U8,U9,U10,U11,U12,
       U13,U14,U15)~1,formula.outcome=Y_obs~1,data=data, nclass=num.tx,
       iter.max=iter,flatten.rhos=1,flatten.gamma=1, stabilize.alphas=1,
       se.method="STANDARD")
  ### Resolve class switching
  cond.mean <- NULL
  # Sum conditional probs (rhos) within each class to determine class names
  cond.mean[1]<-sum(fit.unadj$theta[1:num.item])
  cond.mean[2]<-sum(fit.unadj$theta[(num.item+1):(2*num.item)])
  cond.mean[3]<-sum(fit.unadj$theta[(2*num.item+1):(3*num.item)])
  # Check if classes are well differentiated
  test <- pairwise.diffs(cond.mean)
  if(sum(abs(test)<0.25)>0) print("Warning: Indistinguishable Classes")
  #Reorder classes and summarize lcca
  ord <- order(cond.mean)
  fit.unadj.ord <- permute.class(fit.unadj, ord)

  ### Save Average Treatment Effects, Bias, SE, and MSE
  un.2v1<-summary(fit.unadj.ord)[94][[1]][1]
  un.2v1.b<-summary(fit.unadj.ord)[94][[1]][1]-diff.21
  un.2v1.SE<-summary(fit.unadj.ord)[94][[1]][4]
  un.2v1.MSE <- (un.2v1.b)^2
  un.2v1.cov<-(un.2v1-1.96*un.2v1.SE)<diff.21 & diff.21<(un.2v1+1.96*un.2v1.SE)

  un.3v1<-summary(fit.unadj.ord)[94][[1]][2]
  un.3v1.b<-summary(fit.unadj.ord)[94][[1]][2]-diff.31
  un.3v1.SE<-summary(fit.unadj.ord)[94][[1]][5]
  un.3v1.MSE <- (un.3v1.b)^2
  un.3v1.cov<-(un.3v1-1.96*un.3v1.SE)<diff.31 & diff.31<(un.3v1+1.96*un.3v1.SE)

  un.3v2<-summary(fit.unadj.ord)[94][[1]][3]
  un.3v2.b<-summary(fit.unadj.ord)[94][[1]][3]-diff.32
  un.3v2.SE <-summary(fit.unadj.ord)[94][[1]][6]
  un.3v2.MSE <- (un.3v2.b)^2
  un.3v2.cov<-(un.3v2-1.96*un.3v2.SE)<diff.32 & diff.32<(un.3v2+1.96*un.3v2.SE)

  ### FIT THE LCCA model
  fit.lcca <- lcca(formula.treatment=cbind(U1,U2,U3,U4,U5,U6,U7,U8,U9,
       U10,U11,U12,U13,U14,U15)~X1+X2+X3+X4+X5+X6+X7+X8,
        formula.outcome=Y_obs~X1+X2+X3+X4+X5+X6+X7+X8,data=data,
       nclass=num.tx,iter.max=iter, flatten.rhos=1,flatten.gamma=1,
       stabilize.alphas=1, se.method="STANDARD")
  ### Resolve class switching
  cond.mean <- NULL
  # Sum conditional probs (rhos) within each class to determine class names
  cond.mean[1]<-sum(fit.lcca$theta[1:num.item])
  cond.mean[2]<-sum(fit.lcca$theta[(num.item+1):(2*num.item)])
  cond.mean[3]<-sum(fit.lcca$theta[(2*num.item+1):(3*num.item)])
```

```
  # Check if classes are well differentiated
  test <- pairwise.diffs(cond.mean)
  if(sum(abs(test)<0.25)>0) print("Warning: Indistinguishable Classes")
  #Reorder classes and summarize lcca
  ord <- order(cond.mean)
  fit.lcca.ord <- permute.class(fit.lcca, ord)

  ### Save Average Treatment Effects, Bias, SE, and MSE
  os.2v1<-summary(fit.lcca.ord)[94][[1]][1]
  os.2v1.b<-summary(fit.lcca.ord)[94][[1]][1]-diff.21
  os.2v1.SE<-summary(fit.lcca.ord)[94][[1]][4]
  os.2v1.MSE <- (os.2v1.b)^2
  os.2v1.cov<-(os.2v1-1.96*os.2v1.SE)<diff.21 & diff.21<(os.2v1+1.96*os.2v1.SE)

  os.3v1<-summary(fit.lcca.ord)[94][[1]][2]
  os.3v1.b<-summary(fit.lcca.ord)[94][[1]][2]-diff.31
  os.3v1.SE<-summary(fit.lcca.ord)[94][[1]][5]
  os.3v1.MSE <- (os.3v1.b)^2
  os.3v1.cov<-(os.3v1-1.96*os.3v1.SE)<diff.31 & diff.31<(os.3v1+1.96*os.3v1.SE)

  os.3v2<-summary(fit.lcca.ord)[94][[1]][3]
  os.3v2.b<-summary(fit.lcca.ord)[94][[1]][3]-diff.32
  os.3v2.SE <-summary(fit.lcca.ord)[94][[1]][6]
  os.3v2.MSE <- (os.3v2.b)^2
  os.3v2.cov<-(os.3v2-1.96*os.3v2.SE)<diff.32 & diff.32<(os.3v2+1.96*os.3v2.SE)
  # Save Class Prevalences
  p1<-fit.lcca.ord$marg.gamma[1]
  p2<-fit.lcca.ord$marg.gamma[2]
  p3<-fit.lcca.ord$marg.gamma[3]

  output <- NULL
  output<-as.numeric(c(os.2v1, os.2v1.b, os.2v1.SE, os.2v1.MSE, os.2v1.cov,
                       os.3v1, os.3v1.b, os.3v1.SE, os.3v1.MSE, os.3v1.cov,
                       os.3v2, os.3v2.b, os.3v2.SE, os.3v2.MSE, os.3v2.cov,
                       un.2v1,un.2v1.b,un.2v1.SE,un.2v1.MSE,un.2v1.cov,
                       un.3v1,un.3v1.b,un.3v1.SE,un.3v1.MSE,un.3v1.cov,
                       un.3v2,un.3v2.b,un.3v2.SE,un.3v2.MSE,un.3v2.cov))
  return(output)

}
###############################################################
###############################################################
## 3-step code
###############################################################
###############################################################
threestep.lcca<-function(data, num.tx=3, num.item=15, num.cov=8,
p=c(.33,.33,.34), K=20, iter=100000){

  ### define N = number of individuals
  N <- dim(data)[1]
  ### Calculate true treatment effects
  diff.21 <- mean(data$diff.21)
  diff.31 <- mean(data$diff.31)
  diff.32 <- mean(data$diff.32)

  ###############################################################
  ### LCA Model for Modal & Pseudoclass Methods
  ###############################################################
  fit.lca <- lca(cbind(U1,U2,U3,U4,U5,U6,U7,U8,U9,U10,U11,U12,U13,U14,U15)~1,
       nclass=num.tx,data=data,iter.max=iter,flatten.rho=1,flatten.gammas=1)
  ### Resolve class switching
  cond.mean <- NULL
  # Sum conditional probs (rhos) within each class to determine class names
```

```
cond.mean[1]<-sum(fit.lca$theta[1:num.item])
cond.mean[2]<-sum(fit.lca$theta[(num.item+1):(2*num.item)])
cond.mean[3]<-sum(fit.lca$theta[(2*num.item+1):(3*num.item)])
# Check if classes are well differentiated
test <- pairwise.diffs(cond.mean)
if(sum(abs(test)<0.25)>0) print("Warning: Indistinguishable Classes")
#Reorder classes and summarize lca
ord <- order(cond.mean)
fit.lca.ord <- permute.class(fit.lca, ord)
### SAVE POSTERIOR PROBABILITIES
probs <- fit.lca.ord$post.probs
colnames(probs) <- c("lca.pC1","lca.pC2","lca.pC3")
data <- cbind(data, probs)
################################################################
## CLASS ASSIGNMENT via MODAL ASSIGNMENT CLASS
################################################################
### modal assignment based on largest posterior probability
data$modal_c[data$lca.pC1> data$lca.pC2 & data$lca.pC1> data$lca.pC3] <- 1
data$modal_c[data$lca.pC2> data$lca.pC1 & data$lca.pC2> data$lca.pC3] <- 2
data$modal_c[data$lca.pC3> data$lca.pC1 & data$lca.pC3> data$lca.pC2] <- 3
################################################################
## CLASS ASSIGNMENT via PSEUDOCLASS DRAWS
################################################################
# vector for pseudoclass draws
obs.tx <- matrix(nrow=N, ncol=K)
prob <- cbind(data$lca.pC1, data$lca.pC2, data$lca.pC3)

for(i in 1:N)
{  # set up multinomial RV and take K draws
   # rmultinom generates matrix (dim = num.tx x K)
   T <- rmultinom(K, size = 1, prob=prob[i,])
   # Convert draws matrix to vector, Save observed tx class
   for(k in 1:K)
      {    obs.tx[i,k] <- which.max(T[,k])
      }
}
colnames(obs.tx) <- c("pc1","pc2","pc3","pc4","pc5","pc6","pc7","pc8",
     "pc9","pc10","pc11","pc12","pc13","pc14","pc15","pc16","pc17",
     "pc18","pc19","pc20")
data <- cbind(data,obs.tx)
################################################################
## ESTIMATE TX EFFECTS: UNADJUSTED Modal
################################################################
# outcome regression for Y ~ Tx (Tx is factor variable)
data$modal_c <- factor(data$modal_c, levels=c(1,2,3))
out1un <- lm(Y_obs ~ C(modal_c,contr.treatment(3,base=1)), data=data)
out2un <- lm(Y_obs ~ C(modal_c,contr.treatment(3,base=2)), data=data)
## Class 2 v Class 1
unm.2v1 <- summary(out1un)$coefficient[2,1]
unm.2v1.b <- unm.2v1-diff.21
unm.2v1.SE <- summary(out1un)$coefficient[2,2]
unm.2v1.MSE <- (unm.2v1.b)^2
unm.2v1.cov<-(unm.2v1-1.96*unm.2v1.SE)<diff.21 &
     diff.21<(unm.2v1+1.96*unm.2v1.SE)
## Class 3 v Class 1
unm.3v1 <- summary(out1un)$coefficient[3,1]
unm.3v1.b <- unm.3v1-diff.31
unm.3v1.SE <- summary(out1un)$coefficient[3,2]
unm.3v1.MSE <- (unm.3v1.b)^2
unm.3v1.cov <- (unm.3v1-1.96*unm.3v1.SE)<diff.31 &
     diff.31<(unm.3v1+1.96*unm.3v1.SE)
## Class 3 v Class 2
unm.3v2 <- summary(out2un)$coefficient[3,1]
```

```
  unm.3v2.b <- unm.3v2-diff.32
  unm.3v2.SE <- summary(out2un)$coefficient[3,2]
  unm.3v2.MSE <- (unm.3v2.b)^2
  unm.3v2.cov <- (unm.3v2-1.96*unm.3v2.SE)<diff.32 &
       diff.32<(unm.3v2+1.96*unm.3v2.SE)
  ##############################################################
  ## ESTIMATE PROPENSITY SCORES: Modal assignment
  ##############################################################
  ## modal class Indicators for LATENT CLASS 1, LATENT CLASS 2, LATENT CLASS 3
  data$m1 <-as.numeric(data$modal_c==1)
  data$m2 <-as.numeric(data$modal_c==2)
  data$m3 <-as.numeric(data$modal_c==3)
  ## Run 3 (binary) propensity score models...
  ps.1m <- glm(m1~X1+X2+X3+X4+X5+X6+X7+X8, data=data, family=binomial)
  ps.2m <- glm(m2~X1+X2+X3+X4+X5+X6+X7+X8, data=data, family=binomial)
  ps.3m <- glm(m3~X1+X2+X3+X4+X5+X6+X7+X8, data=data, family=binomial)
  ## Extract estimated PS from binary PS models
  ps1m <- predict(ps.1m, type = "response")
  ps2m <- predict(ps.2m, type = "response")
  ps3m <- predict(ps.3m, type = "response")
  ## Calculate ATE weights from binary PS models
  w.1m <- ifelse(data$modal_c==1, 1/ps1m, 1/(1-ps1m))
  w.2m <- ifelse(data$modal_c==2, 1/ps2m, 1/(1-ps2m))
  w.3m <- ifelse(data$modal_c==3, 1/ps3m, 1/(1-ps3m))
  ## Set final estimated ATE weight to ATE weight calc from observed LC
  data$w.fm[data$modal_c==1] <- w.1m[data$modal_c==1]
  data$w.fm[data$modal_c==2] <- w.2m[data$modal_c==2]
  data$w.fm[data$modal_c==3] <- w.3m[data$modal_c==3]
  ## Set weights above 98th percentile to 98th percentile value
  data$w.fm[data$w.fm>quantile(data$w.fm,probs=.98)]<-
       quantile(data$w.fm,probs=.98)
##############################################################
## ESTIMATE PROPENSITY SCORES: Pseudoclass assignment
##############################################################
# Matrix for final pseudoclass weights
w.fp <- pc.w <- NULL
pc_names <- c("pc1","pc2","pc3","pc4","pc5","pc6","pc7","pc8","pc9","pc10",
       "pc11","pc12","pc13","pc14","pc15","pc16","pc17","pc18","pc19","pc20")
for(k in 1:K)
{       kk <- pc_names[k]
   ## Indicators for LATENT CLASS 1, LATENT CLASS 2, LATENT CLASS 3
   pc1.ind <- as.numeric(data[,kk]==1)
   pc2.ind <- as.numeric(data[,kk]==2)
   pc3.ind <- as.numeric(data[,kk]==3)
   ## Run 3 (binary) propensity score models...
   ps.1p <-glm(pc1.ind~X1+X2+X3+X4+X5+X6+X7+X8, data = data, family = binomial)
   ps.2p <-glm(pc2.ind~X1+X2+X3+X4+X5+X6+X7+X8, data = data, family = binomial)
   ps.3p <-glm(pc3.ind~X1+X2+X3+X4+X5+X6+X7+X8, data = data, family = binomial)
   ## Extract estimated PS from binary PS models
   ps1p <- predict(ps.1p, type = "response")
   ps2p <- predict(ps.2p, type = "response")
   ps3p <- predict(ps.3p, type = "response")
   ## Calculate ATE weights from binary PS models
   w.1p <- ifelse(pc1.ind==1, 1/ps1p, 1/(1-ps1p))
   w.2p <- ifelse(pc2.ind==1, 1/ps2p, 1/(1-ps2p))
   w.3p <- ifelse(pc3.ind==1, 1/ps3p, 1/(1-ps3p))
   ## Set final estimated ATE weight to ATE weight calc from observed LC
   w.fp[pc1.ind==1] <- w.1p[pc1.ind==1]
   w.fp[pc2.ind==1] <- w.2p[pc2.ind==1]
   w.fp[pc3.ind==1] <- w.3p[pc3.ind==1]
   ## Set weights above 98th percentile to 98th percentile value
   w.fp[w.fp>quantile(w.fp,probs=.98)] <- quantile(w.fp,probs=.98)
   pc.w <- cbind(pc.w,w.fp)
```

```
}
colnames(pc.w) <- c("w1p","w2p","w3p","w4p","w5p","w6p","w7p","w8p","w9p",
"w10p","w11p","w12p","w13p","w14p","w15p","w16p","w17p","w18p","w19p","w20p")
data <- cbind(data, pc.w)
 ################################################################
  ## ESTIMATE TX EFFECTS: Modal + Propensity Scores
  ################################################################
  m.design <- svydesign(id=~1, weights=~w.fm, data=data)
  # outcome regression for Y ~ Tx (Tx is factor variable)
  data$modal_c <- factor(data$modal_c, levels=c(1,2,3))
  out1m<-svyglm(Y_obs~C(modal_c,contr.treatment(3,base=1)),
       design=m.design,data=data)
  out2m <- svyglm(Y_obs ~ C(modal_c,contr.treatment(3,base=2)),
       design=m.design,data=data)
  ## Class 2 v Class 1
  m.2v1 <- summary(out1m)$coefficient[2,1]
  m.2v1.b <- m.2v1-diff.21
  m.2v1.SE <- summary(out1m)$coefficient[2,2]
  m.2v1.MSE <- (m.2v1.b)^2
  m.2v1.cov <- (m.2v1-1.96*m.2v1.SE)<diff.21 & diff.21<(m.2v1+1.96*m.2v1.SE)
  ## Class 3 v Class 1
  m.3v1 <- summary(out1m)$coefficient[3,1]
  m.3v1.b <- m.3v1-diff.31
  m.3v1.SE <- summary(out1m)$coefficient[3,2]
  m.3v1.MSE <- (m.3v1.b)^2
  m.3v1.cov <- (m.3v1-1.96*m.3v1.SE)<diff.31 & diff.31<(m.3v1+1.96*m.3v1.SE)
  ## Class 3 v Class 2
  m.3v2 <- summary(out2m)$coefficient[3,1]
  m.3v2.b <- m.3v2-diff.32
  m.3v2.SE <- summary(out2m)$coefficient[3,2]
  m.3v2.MSE <- (m.3v2.b)^2
  m.3v2.cov <- (m.3v2-1.96*m.3v2.SE)<diff.32 & diff.32<(m.3v2+1.96*m.3v2.SE)
  ################################################################
  ## ESTIMATE TX EFFECTS: Pseudoclass + Propensity Scores
  ################################################################
  # Matrix for estimated treatment effects
  q.matrix <- matrix(nrow=K, ncol=choose(num.tx,2))
  se.matrix <- matrix(nrow=K, ncol= choose(num.tx,2))
  pcw_names<- c("w1p","w2p","w3p","w4p","w5p","w6p","w7p","w8p","w9p","w10p",
        "w11p","w12p","w13p","w14p","w15p","w16p","w17p","w18p","w19p","w20p")
  ### For each pseudoclass draw, estimate Tx Effect and SE
  for(k in 1:K)
  {     # Define temporary variable: kth pseudoclass Tx Class
     data$tx_temp <- factor(data[,pc_names[k]],levels = c(1,2,3))
     # Define temporary variable: kth pseudoclass PS Weight
     data$w_temp <- data[,pcw_names[k]]
     # define survey design
     p.design<- svydesign(id=~1, weights=~w_temp, data=data)
     # PS weighted outcome regression for Y ~ Tx
     out1p <- svyglm(Y_obs ~ C(tx_temp,contr.treatment(3,base=1)),
       design=p.design, data=data)
     out2p <- svyglm(Y_obs ~ C(tx_temp,contr.treatment(3,base=2)),
       design=p.design, data=data)

     ## Class 2 v Class 1 contrast, SE
     q.matrix[k,1] <- summary(out1p)$coefficient[2,1]
     se.matrix[k,1] <- summary(out1p)$coefficient[2,2]
     ## Class 3 v Class 1 contrast, SE
     q.matrix[k,2] <- summary(out1p)$coefficient[3,1]
     se.matrix[k,2] <- summary(out1p)$coefficient[3,2]
     ## Class 3 v Class 2 contrast, SE
     q.matrix[k,3] <- summary(out2p)$coefficient[3,1]
     se.matrix[k,3] <- summary(out2p)$coefficient[3,2]
```

```
  }

  ### Combine Tx Effects, SE across PC. Then calc Bias, MSE, 95% CI coverage
  p.2v1 <- as.numeric(MIcombine(as.list(q.matrix[,1]),
       as.list(se.matrix[,1]))[1])
  p.2v1.b <- p.2v1 - diff.21
  p.2v1.SE <- as.numeric(MIcombine(as.list(q.matrix[,1]),
       as.list(se.matrix[,1]))[2])
  p.2v1.MSE <- (p.2v1.b)^2
  p.2v1.cov <- (p.2v1-1.96*p.2v1.SE)<diff.21 & diff.21<(p.2v1+1.96*p.2v1.SE)

  p.3v1 <- as.numeric(MIcombine(as.list(q.matrix[,2]),
       as.list(se.matrix[,2]))[1])
  p.3v1.b <- p.3v1 - diff.31
  p.3v1.SE <- as.numeric(MIcombine(as.list(q.matrix[,2]),
       as.list(se.matrix[,2]))[2])
  p.3v1.MSE <- (p.3v1.b)^2
  p.3v1.cov <- (p.3v1-1.96*p.3v1.SE)<diff.31 & diff.31<(p.3v1+1.96*p.3v1.SE)

  p.3v2 <- as.numeric(MIcombine(as.list(q.matrix[,3]),
       as.list(se.matrix[,3]))[1])
  p.3v2.b <- p.3v2 - diff.32
  p.3v2.SE <- as.numeric(MIcombine(as.list(q.matrix[,3]),
       as.list(se.matrix[,3]))[2])
  p.3v2.MSE <- (p.3v2.b)^2
  p.3v2.cov <- (p.3v2-1.96*p.3v2.SE)<diff.32 & diff.32<(p.3v2+1.96*p.3v2.SE)
  ### Class prevalences from LCA (same for Modal and Pseudoclass)
  p1<-summary(fit.lca.ord)[66][[1]][1]
  p2<-summary(fit.lca.ord)[66][[1]][2]
  p3<-summary(fit.lca.ord)[66][[1]][3]

  output <- NULL
  output<-as.numeric(c(unm.2v1,unm.2v1.b,unm.2v1.SE,unm.2v1.MSE,unm.2v1.cov,
                      unm.3v1,unm.3v1.b,unm.3v1.SE,unm.3v1.MSE,unm.3v1.cov,
                      unm.3v2,unm.3v2.b,unm.3v2.SE,unm.3v2.MSE,unm.3v2.cov,
                      m.2v1,m.2v1.b,m.2v1.SE,m.2v1.MSE,m.2v1.cov,
                      m.3v1,m.3v1.b,m.3v1.SE,m.3v1.MSE,m.3v1.cov,
                      m.3v2,m.3v2.b,m.3v2.SE,m.3v2.MSE,m.3v2.cov,
                      p.2v1,p.2v1.b,p.2v1.SE,p.2v1.MSE,p.2v1.cov,
                      p.3v1,p.3v1.b,p.3v1.SE,p.3v1.MSE,p.3v1.cov,
                      p.3v2,p.3v2.b,p.3v2.SE,p.3v2.MSE,p.3v2.cov))
  return(output)

}
################################################################
################################################################
## PAIRWISE Diff function
################################################################
################################################################
## CALCULATE PAIRWISE DIFFS
## Solve class switching by comparing sum of estimated item prob
## across classes. Since data is set up with monotonically increasing
## item probs across classes, sum(C1 item) < sum(C2 items) < sum(C3 items)
## This function compares pairwise differences of these sums to
## ensure that they are sufficiently large: i.e. classes are
## adequately differentiated
################################################################
pairwise.diffs <- function(x)
{   # create combination pairs
  prs <- cbind(rep(1:length(x), each = length(x)), 1:length(x))
  # drop ones that compare same classes
  drops <- NULL
  for(i in 1:(length(x))^2 )
```

```
  { if (prs[i,1]==prs[i,2]) drops <- c(drops, i)
  }
  new <- prs[-drops,]
  # do pairwise differences
  result <- x[new[,1]] - x[new[,2], drop = FALSE]
}
################################################################
```

CURRICULUM VITAE

MEGAN S. SCHULER

Department of Mental Health mschuler@jhsph.edu
Johns Hopkins Bloomberg School of Public Health (803) 730-3476
624 N Broadway HH850
Baltimore MD 21205

EDUCATION

2013 **Johns Hopkins Bloomberg School of Public Health**
Ph.D. in Mental Health
Advisor: Elizabeth Stuart
Dissertation title: *Estimating the relative treatment effects of natural clusters of adolescent substance abuse treatment services: Combining latent class analysis and propensity score methods*

2009 **Medical University of South Carolina**
M.S. in Biostatistics
Advisors: Rickey Carter, Stacia DeSantis
Thesis title: *Concordance of urine drug screens and self-reported use in cocaine clinical trials*

2006 **Tulane University**
B.S. with honors in Mathematics
Honors thesis title: *A rational Landen transformation of degree eight*
Summa Cum Laude, Phi Beta Kappa Society

2005 **Budapest Semesters in Mathematics**

PROFESSIONAL EXPERIENCE

Pennsylvania State University, State College PA starting Jan 2014
Postdoctoral Fellow
National Institute on Drug Abuse Prevention and Methodology Training Program
The Prevention Research Center and The Methodology Center

RAND Corporation, Arlington VA Summer 2011
Summer Associate
Worked with the Drug Policy Research Group under the guidance of Dr. Beth Ann Griffin to investigate the effectiveness of adolescent substance abuse treatment programs. Research topics included propensity score analysis with multiple treatment groups and latent class modeling. Designed and coded R and Mplus programs.

Drug Policy Alliance, Washington DC Fall 2010
Intern

Performed literature reviews and wrote policy briefs relating to evidence-based public health interventions for substance use.

Center for Drug & Alcohol Problems, Medical University of South Carolina
Research Assistant 2007-2009
Served as statistician for Dr. Robert Malcolm's substance abuse research group. Gained experience in the conduct and analysis of clinical trials for cocaine dependence. Provided statistical analysis, database management, and Data and Safety Monitoring Board reports. Designed and coded SAS programs.

Baruch Institute for Marine & Coastal Sciences, University of South Carolina
Lab Manager, Research Assistant 2006-2007
Performed analysis on physical oceanographic data, designed and coded MATLAB programs, and produced technical reports. Maintained lab instruments and webpage.

Tulane University, New Orleans LA Summer 2005
Research Assistant
Performed individual research under the guidance of Dr. Victor Moll relating to dynamical systems arising from Landen transformation integrals. Designed and coded Mathematica programs.

Rice University Summer Institute of Statistics, Houston TX Summer 2004
Student Researcher
Research project focused on data imputation techniques for large-scale microarray data. Performed statistical analysis, designed and coded R programs.

AWARDS

2013	Morton Kramer Annual Award (*JHSPH departmental award for biostatistical achievement*)
2013	ASA Health Policy Statistics Section student paper award
2012	Selected to present at Student Research Showcase at National Conference on Health Statistics
2012, 2013	Society for Prevention Research conference Student Travel Award
2011	International Conference of Health Policy Statistics Student Travel Award
2009-2012	Alfred Sommer Merit Scholarship (*full tuition*) (JHSPH)
2008-2009	Presidential Scholar (MUSC)
2008-2009	GAANN Fellowship in New Technologies and Human Biology (MUSC)
2007-2008	Dean's Honor Scholarship (*full tuition*) (MUSC)
2006	Senior Scholar in Mathematics (Tulane)
2006	Departmental Honors (Tulane)
2006	Kappa Kappa Gamma Prize in Mathematics (Tulane)
2006	Glendy Burke Medal in Mathematics (Tulane)
2002-2006	Dean's Honor Scholarship (*full tuition*) (Tulane)

CONFERENCE PRESENTATIONS

Invited Presentations

2013 International Conference on Health Policy Statistics (ICHPS), Chicago IL
Estimating causal effects of latent treatment classes: Natural clusters of drug treatment services for adolescents

2013 Joint Statistical Meeting (JSM), Montreal Canada
Effectiveness of biological drug testing among adolescent substance users: A multiple group propensity score analysis

2013 Society of Prevention Research (SPR) conference, San Francisco CA
Estimating causal effects when treatment is modeled as a latent variable and an application to adolescent drug treatment

2012 National Conference on Health Statistics (NCHS), Washington DC
Generalizing observational study results: Applying propensity score methods to complex surveys

Contributed Oral Presentations

2012 Joint Statistical Meeting (JSM), San Diego CA
Modeling naturalistic clusters of adolescent substance abuse treatment services: A latent class analysis

2008 MUSC Student Research Day, Charleston SC
Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies

2006 American Geophysical Union (AGU), San Francisco CA
Application of a high-frequency radar to an inter-tidal salt marsh

Poster Presentations

2013 AcademyHealth's Annual Research Meeting, Baltimore MD
Generalizing observational study results: Applying propensity score methods to complex surveys

2012 College on Problems of Drug Dependence (CPDD), Palm Springs CA
Effectiveness of adolescent substance abuse treatments: Is drug screening sufficient?

2012 Society for Prevention Research (SPR) conference, Washington DC
Modeling substance abuse services received by adolescents as latent classes

2012 Atlantic Causal Inference Conference (ACIC), Baltimore MD
Using generalized boosted models for propensity score estimation for multiple treatments: A substance abuse treatment application

2011 International Conference on Health Policy Statistics (ICHPS), Cleveland OH
Effectiveness of adolescent substance abuse treatments: An application using multinomial propensity scores

2009 College on Problems of Drug Dependence (CPDD), Reno NV
Impulsivity, completion status, and ethnicity in a clinical trial for cocaine dependence

2008 Society for Advancement of Chicanos & Native Americans in Science (SACNAS) National Conference, Salt Lake City UT
Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies

PUBLICATIONS

Schuler, M. S., Griffin, B. A., Ramchand, R., Almirall, D., & McCaffrey, D. Effectiveness of treatment for adolescent substance use: Is biological drug testing sufficient? *Journal of Studies on Alcohol and Drugs*, [in press].

DuGoff, E., Schuler, M. S., & Stuart, E. A. (2013). Generalizing propensity score results: Applying matching methods to complex surveys. Forthcoming in *Health Services Research*. Published online 16 June 2013. DOI: 10.1111/1475-6773.12090.

Schuler, M. S., Lechner, W. V., Carter, R., & Malcolm, R. (2009). Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies. *Journal of Addiction Medicine*, 3(4): 211-217. DOI: 10.1097/ADM.0b013e3181a0f5dc. http://www.ncbi.nlm.nih.gov/pubmed/20209029

Manuscripts Under Review And In Preparation

Schuler, M. S., Griffin, B. A., Letourneau, E., & Stuart, E. A. Common clusters of outpatient drug treatment services for adolescents: A latent class analysis. [Under review, *Journal of Substance Abuse Treatment*].

Schuler, M. S., & Stuart, E. A. One-step and three-step methods for categorical latent variable regression. [In preparation for *Psychological Methods*].

Schuler, M. S., Puttaiah, S., Motjabai, R., & Crum, R. Barriers to accessing treatment for alcohol problems: A latent class analysis. [In preparation for *Psychiatric Services*].

Levin, S., Kirsch, T., Dugas, A., Yenokyan, G., Schuler, M. S., Toerper, M., Korley, F., France, D., Hager, D., & Anderson, G. Hospital access and insurance for emergency department patients in the United States. [In preparation for *JAMA*].

Darnell, A. J., Herz, D., Schuler, M. S., & Pecora, P. J. Quasi-experimental evaluation of Functional Family Therapy effectiveness with a diverse urban sample. [In preparation for *Children and Youth Services Review*].

Book Chapters

Schuler, M. S. A review of computer programs for propensity score analysis. In Wan Fei, Haiyan Bai, eds. *Propensity Score Analysis: Fundamentals, Developments, and Extensions.* New York: Guilford Press. [In press]

TEACHING EXPERIENCE

Teaching Assistant

2012-2013	Department of Mental Health, JHSPH Class: Causal Inference in Medicine and Public Health
2011-2012	Department of Mental Health, JHSPH Class: Psychosocial Methods I: Measurement
2011-2012	Department of Mental Health, JHSPH Class: Psychosocial Methods II: Structural Equation Modeling
2011	Department of Mental Health, JHSPH Class: Drugs, Society & Policy
2009	Department of Biostatistics, Epidemiology, and Bioinformatics; MUSC Class: Regression Methods in Clinical Research
2006	Department of Mathematics, Tulane University Class: Introduction to Probability & Statistics

Instructor

2006-2007 Kaplan Test Prep

Training

2009 Teaching Techniques class (MUSC)

PROFESSIONAL SERVICE

American Statistical Association

- Session Chair, Joint Statistical Meetings 2012
- Member 2011-present

College of Problems on Drug Dependence Member 2012-present

Society for Prevention Research Member 2012-present

Johns Hopkins Bloomberg School of Public Health, Department of Mental Health

- President of the Mental Health Student Group 2011-2012
- Vice President of the Mental Health Student Group 2012-2013

Ad-hoc reviewer

- *Evaluation and Program Planning*
- *Health Services and Outcomes Research Methodology*

CONSULTING EXPERIENCE

University Research Co. LLC, Bethesda MD 2013-present

Provided statistical guidance regarding missing data.

Casey Family Programs, Seattle WA 2012-2013

Provided propensity score analysis and statistical guidance.

Resource Associates Inc., Columbia SC 2006-2011

Provided statistical analysis and grant writing assistance related to program evaluation of school-based educational interventions.